

09/117246

# BREVET D'INVENTION

**CERTIFICAT D'UTILITÉ - CERTIFICAT D'ADDITION**

## COPIE OFFICIELLE

Le Directeur général de l'Institut national de la propriété industrielle certifie que le document ci-annexé est la copie certifiée conforme d'une demande de titre de propriété industrielle déposée à l'Institut.

Fait à Paris, le

10 OCT. 1997

Pour le Directeur général de l'Institut  
national de la propriété industrielle  
Le Chef de Division

**PRIORITY DOCUMENT**

Yves CAMPENON

INSTITUT  
NATIONAL DE  
LA PROPRIÉTÉ  
INDUSTRIELLE

SIEGE  
26 bis, rue de Saint Petersburg  
75800 PARIS Cedex 08  
Téléphone : 01 53 04 53 04  
Télécopie : 01 42 93 59 30



**REQUETE**  
**EN DÉLIVRANCE D'UN**  
**TITRE DE PROPRIÉTÉ**  
**INDUSTRIELLE \***

1

a	<input checked="" type="checkbox"/>	BREVET D'INVENTION
b	<input type="checkbox"/>	CERTIFICAT D'UTILITÉ
c	<input type="checkbox"/>	DEMANDE DIVISIONNAIRE
d	<input type="checkbox"/>	TRANSFORMATION D'UNE DEMANDE DE BREVET EUROPEEN

Pour c et d, précisez : Nature, N° et date de la demande initiale

**2 OPTIONS OBLIGATOIRES** au moment du dépôt (sauf pour le certificat d'utilité)

LE DEMANDEUR REQUIERT L'ÉTABLISSEMENT DIFFERÉ DU RAPPORT DE RECHERCHE \*

☐ OUI

☒ NON

SI L'OPTION CHOISIE EST NON ET SI LE DEMANDEUR EST UNE PERSONNE PHYSIQUE IL REQUIERT LE PAIEMENT ÉCHELONNÉ DE LA REDEVANCE DE RAPPORT DE RECHERCHE

☐ OUI

☐ NON

NATURE

NUMERO

DATE DE LA DEMANDE INITIALE

DATE DE REMISE DES PIÈCES

29 JAN 1996

N° D'ENREGISTREMENT NATIONAL

96 01004 -

DATE DE DÉPÔT

29 JAN 1996

CODE POSTAL DU LIEU DE DÉPÔT

4 NUMERO DU POUVOIR PERMANENT

**3 NOM ET ADRESSE DU DEMANDEUR OU DU MANDATAIRE A QUI TOUTE LA CORRESPONDANCE DOIT ÊTRE ADRESSÉE**

ERNEST GUTMANN-  
YVES PLASSERAUD S.A.  
3 rue Chauveau-Lagarde  
75008 PARIS (France).

5 REFERENCE DU CORRESPONDANT

B2928 - AD

6 TELEPHONE DU CORRESPONDANT

(1) 44 51 18 00

**7 TITRE DE L'INVENTION**

PROTEINES DE RESERVE DE PLANTES ENRICHIES EN ACIDES AMINES,  
NOTAMMENT  $\gamma$ -ZEINE DE MAIS ENRICHIE EN LYSINE. PLANTES EXPRIMANT  
CES PROTEINES.

**8 DEMANDEUR(S) : Nom et Prénoms (souligner le nom patronymique) ou dénomination et forme juridique**

N° SIREN.

BIOCEM

Société anonyme

**9 ADRESSE(S) COMPLÈTE(S)**

Campus Universitaire des Cézeaux  
24 avenue des Landais  
63170 AUBIERE

PAYS

FRANCE

**10 NATIONALITÉ(S)**

Française

**11 INVENTEUR(S)**

LE DEMANDEUR EST L'UNIQUE INVENTEUR \*

☐ OUI

Si la réponse est non voir notice explicative

☒ NON

**12**

SI LE DEMANDEUR EST UNE PERSONNE PHYSIQUE NON IMPOSABLE IL REQUIERT OU A REQUIS LA RÉDUCTION DES REDEVANCES \*

☐ OUI

☐ NON

DE DÉPÔT

REDEVANCES VERSÉES

DE RAPPORT DE RECHERCHE

DE REVENDEMENT DE PRIORITÉ

DE REVENDEMENT (à partir de la 11e)

**13 DÉCLARATION DE PRIORITÉ**

DU REQUÊTE DU BÉNÉFICE DE LA DATE DE DÉPÔT D'UNE DEMANDE ANTERIEURE

PAYS D'ORIGINE

DATE DE DÉPÔT

NUMERO

**14**

DIVISIONS

ANTÉRIEURES A LA PRÉSENTE DEMANDE

N°

N°

N°

N°

**15 SIGNATURE DU DEMANDEUR OU DU MANDATAIRE**  
NOM ET QUALITÉ DU SIGNATAIRE-N° D'INSCRIPTION

Anne DESATX  
CIP N° 933006

**SIGNATURE DU PRÉPOSÉ A LA RÉCEPTION**

**SIGNATURE APRES ENREGISTREMENT DE LA DEMANDE A L'INPI**

## Division Administrative des Brevets

## DÉSIGNATION DE L'INVENTEUR

(si le demandeur n'est pas l'inventeur ou l'unique inventeur)

N° d'enregistrement national

9601004

## Titre de l'invention :

PROTEINES DE RESERVE DE PLANTES ENRICHIES EN ACIDES AMINES,  
NOTAMMENT  $\gamma$ -ZEINE DE MAIS ENRICHIE EN LYSINE, PLANTES EXPRIMANT  
CES PROTEINES.

## Le (s) soussigné (s)

ERNEST GUTMANN-  
YVES PLASSERAUD S.A.  
3 rue Chauveau-Lagarde  
75008 PARIS (France).

désigne (nt) en tant qu'inventeur (s) (indiquer nom, prénoms, adresse et souligner le nom  
patronymique) :

1) LUDEVID Dolorès  
CENTRO DE INVESTIGACION Y DESARROLO. CSIC  
Jordi Girona Salgado 18-26  
08034 BARCELONE (Espagne)

2) TORRENT Margarita  
CENTRO DE INVESTIGACION Y DESARROLLO. CSIC  
Jordi Girona Salgado 18-26  
08034 BARCELONE (Espagne)

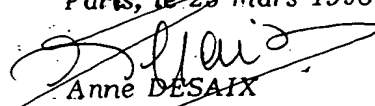
3) ALVAREZ Inaki  
CENTRO DE INVESTIGACION Y DESARROLO. CSIC  
Jordi Girona Salgado 18-26  
08034 BARCELONE (Espagne)

4) PEREZ Pascual  
21 allée du Parc  
63110 BEAUMONT (France)

**NOTA :** A titre exceptionnel, le nom de l'inventeur peut être suivi de celui de la société à laquelle il appartient  
(société d'appartenance) lorsque celle-ci est différente de la société déposante ou titulaire.

Date et signature (s) du (des) demandeur (s) ou du mandataire

Paris, le 25 mars 1996

  
Anne DESAIX  
CPI N° 93.3006

Protéines de réserve de plantes enrichies en acides aminés, notamment  $\gamma$ -zéine de maïs enrichie en lysine. Plantes exprimant ces protéines

5 La présente demande a pour objet de nouveaux moyens permettant la préparation de plantes exprimant des protéines de réserve enrichies en acides aminés déficitaires dans les protéines de réserve normales, et en particulier de protéines de réserves enrichies en lysine. L'invention a également pour objet les protéines de réserve ainsi modifiées, ainsi que  
10 des plantes exprimant ces protéines de réserve modifiées.

De nombreuses plantes, le cas échéant après transformation par des étapes physico-chimiques, sont d'une importance économique majeure dans l'alimentation humaine ou animale, et le problème de l'amélioration de leur qualité nutritionnelle a déjà donné lieu à différents types de  
15 recherches. En particulier, pour remédier à l'insuffisante présence de certains acides aminés dans les protéines de réserve des plantes, on a mis au point des variétés sélectionnées présentant des qualités nutritionnelles supérieures, ou proposé différentes modifications faisant appel à l'utilisation des techniques du génie génétique afin de favoriser ou  
20 d'augmenter la production dans ces plantes notamment de certains acides aminés déficitaires, et néanmoins importants pour les qualités nutritionnelles de la plante. Ces acides aminés déficitaires sont par exemple la lysine, la méthionine.

Dans le cadre de la présente demande, les inventeurs se sont  
25 proposé d'apporter une solution originale au problème de l'amélioration des plantes, notamment de l'amélioration de leurs qualités nutritionnelles et ont à cet effet travaillé en premier lieu à partir d'une plante d'importance économique considérable, le maïs. Plus précisément, ils se sont intéressés  
aux protéines de réserve de l'endosperme des semences de maïs,

protéines qui comprennent notamment les zéines et, en particulier, la  $\gamma$ -zéine.

Au cours du développement des semences de maïs, les cellules de l'endosperme synthétisent de grandes quantités de protéines de réserve, et notamment de  $\alpha$ -,  $\beta$ -, et  $\gamma$ -zéines. Ces zéines sont accumulées dans les corps protéiques dérivés du réticulum endoplasmique (ER).

D'une façon générale, les zéines représentent un groupe complexe de protéines, réparties en plusieurs groupes,  $\alpha$ -,  $\beta$ -,  $\gamma$ -, et  $\delta$ -zéines (Larkins et al, 1989), codées par une famille génique (Hagen et Rubenstein, 1980, Gene 13, 239-249). Bien qu'elles présentent une structure variable, ces protéines ont des caractéristiques communes : la présence dans leur structure primaire de répétitions en tandem riches en résidus d'acides aminés de type Proline, la présence de nombreux résidus hydrophobes qui confèrent leur insolubilité à ces protéines dans un milieu aqueux, et l'absence de résidus lysine, acides aminés essentiels pour l'homme et les animaux monogastriques. L'absence de lysine dans toutes les protéines majoritaires (détectées en grande quantité dans l'endosperme) naturellement produites de la famille des zéines conduit à une composition en acides aminés déséquilibrée dans les semences de maïs.

Parmi ces protéines, la  $\gamma$ -zéine de maïs est une protéine ayant un poids moléculaire de 28 kDa, dont la séquence codante a été décrite sous la forme de l'ADNc par Prat et al (Nucleic Acids Research, vol. 13, n° 5, 1985, p. 1494-1504). La séquence complète du gène codant pour la  $\gamma$ -zéine, y compris les séquences non codantes en amont et en aval contenant les éléments de régulation de l'expression, a été décrite par Reina, M. et al (Nucleic Acids Research, vol. 18, n° 21, 1990, p 6426).

Jusqu'à présent, différentes approches ont été envisagées pour augmenter la teneur en lysine des protéines de la famille des zéines. A cet égard, des approches génétiques et moléculaires ont été mises en oeuvre. Par exemple, des mutants permettant l'obtention de maïs riche en lysine,

tels que le mutant opaque-2 (o2) et le mutant floury-2 (fl-2) (Mertz et al, 1964, Science 145, 279-280, Nelson et al, 1965, Science 150, 1469-1470) ont ainsi été proposés et des tentatives pour remédier aux effets néfastes de l'absence de certaines classes de zéïnes, notamment les  $\alpha$ -zéïnes, sur les caractéristiques phénotypiques, par sélection de maïs contenant des gènes modificateurs o2 (Paez et al, 1969, Plant Sci. 9, 251-252, Geetha et al, 1991, Plant Cell 3, 1207-1219) ont été faites.

Une autre approche a consisté à réaliser une action indirecte sur la production de lysine libre, en particulier dans les plantes dicotylédones. Cette technique a été mise en oeuvre en agissant sur la dérégulation des enzymes clé (DHAPS et AK) impliquées dans le cycle de la biosynthèse de la lysine via l'aspartate. Un accroissement sensible des niveaux de lysine libre a été obtenu dans les feuilles, mais pas dans les semences, dans des expériences de transformation de plantes de tabac avec des bactéries E. coli contenant les gènes dapA et des bactéries E. coli contenant le gène lysC (Shaul et Galili, 1992, Plant J. 2, 203-209 et 1993, Plant Mol. Biol. 23, 759-768 ; Perl, A., Schaul O., Galili, G., 1992, Plant Molecular Biology, 19, p. 815-823). Récemment, les mêmes gènes dapA de Corynebacterium et lysC de E. coli ont été utilisés et exprimés sous le contrôle d'un promoteur spécifique des semences dans des plantes de soja. L'expression de ces deux enzymes dans le soja a conduit à une augmentation de cinq fois de la teneur en lysine dans les semences (Falco et al, 1995, BIO-Technology 13, 577-582).

D'autres auteurs (Wallace et al, 1988, Science 240, 662-664) ont tenté d'augmenter la lysine dans l' $\alpha$ -zéïne (19 kDa) des semences de maïs en incorporant de façon ponctuelle des résidus lysine en différentes positions dans la molécule d' $\alpha$ -zéïne. L'expression de ces constructions dans des oocytes de Xenopus a conduit à l'assemblage correct des zéïnes riches en lysine dans les vésicules analogues des corps protéiques. Cependant, l' $\alpha$ -zéïne normale et la forme modifiée enrichie en lysine

étaient dégradées lorsqu'elles étaient exprimées dans des semences de tabac (Othani et al, 1991, Plant Mol. Biol. 16:117).

On ne disposait donc pas à ce jour de connaissances quant aux moyens pouvant permettre l'expression d'une zéine enrichie en lysine dans les cellules la produisant naturellement dans le maïs, c'est-à-dire dans les cellules de l'endosperme. A fortiori l'expression dans d'autres cellules végétales de zéines enrichies en lysine n'était pas maîtrisée.

L'un des buts de l'invention est donc de proposer des moyens permettant d'obtenir une zéine enrichie en lysine, en particulier une  $\gamma$ -zéine de maïs enrichie en lysine, cette protéine étant exprimée notamment dans les cellules des semences de maïs et en particulier dans les cellules de l'endosperme, ladite protéine modifiée étant en outre exprimée de façon telle que ses propriétés quant à la localisation et l'accumulation au niveau du réticulum endoplasmique et des corps protéiques dérivés soient préservées.

L'expression « enrichie en lysine » signifie, dans le contexte de la présente demande, que la protéine comporte un nombre augmenté de résidus lysine par rapport à la protéine naturelle dont elle est dérivée, par exemple par suite d'une modification de la séquence de nucléotides qui l'exprime.

L'invention propose aussi des moyens pour obtenir l'expression de protéines, de préférence de  $\gamma$ -zéines enrichies en lysine, dans des cellules végétales de différents tissus, par exemples des tissus de feuilles ou de racines, et ce le cas échéant dans des cellules végétales de plantes n'exprimant pas naturellement la protéine, en particulier la  $\gamma$ -zéine dont la production est recherchée.

Par ailleurs, selon un mode de réalisation particulier de l'invention, d'autres protéines de réserve peuvent être enrichies en lysine selon des modalités analogues.

Les inventeurs se sont proposé dans un premier temps, pour réaliser l'invention, d'introduire dans le gène codant pour la  $\gamma$ -zéine ou pour d'autres protéines de réserve du maïs ou d'autres plantes, ou dans la séquence codante de ce gène, des séquences codant pour des polypeptides enrichis en lysine, afin d'obtenir la production de  $\gamma$ -zéines ou d'autres protéines enrichies en lysine et donc, de semences enrichies en lysine. Différents sites de la séquence codante du gène de la  $\gamma$ -zéine ont été identifiés comme sites permissifs, (appelés encore "sites neutres") pour préparer des séquences de nucléotides ainsi modifiées.

La présente demande propose donc des moyens pour transformer le gène codant pour la  $\gamma$ -zéine de maïs ou pour transformer toute séquence de nucléotides codant pour la  $\gamma$ -zéine et dérivée de ce gène, de façon à obtenir à partir de l'expression du gène modifié ou plus généralement de la séquence de nucléotides modifiée, une protéine enrichie en lysine; ces moyens comprennent en particulier des oligonucléotides synthétiques codant pour une séquence d'acides aminés comportant des résidus lysine.

L'invention a également pour objet des séquences de nucléotides recombinantes ou séquences chimères susceptibles de coder pour une  $\gamma$ -zéine enrichie en lysine.

Entrent également dans le cadre de l'invention des cellules hôtes transformées par de telles séquences, et en particulier des cellules végétales, par exemple des cellules permettant la régénération de plantes, ainsi que des plantes ou des parties de plantes (tissus, organes,...) contenant de telles cellules et produisant de façon stable des protéines de réserve modifiées, et en particulier des  $\gamma$ -zéines enrichies en lysine.

L'invention a également pour objet lesdites protéines modifiées, par exemple enrichies en lysine, ainsi que des anticorps dirigés contre ces protéines.

Un oligonucléotide approprié pour la réalisation de l'invention, utilisable pour la préparation de séquences de nucléotides recombinantes,

est caractérisé en ce qu'il comprend au moins un enchaînement codant pour un polypeptide répondant à la formule  $(P-K)_n$ , dans laquelle :

- n est un entier supérieur ou égal à 2,
- P représente un résidu d'acide aminé Proline,
- K représente un résidu d'acide aminé Lysine,
- le signe "-" symbolise une liaison entre les deux résidus d'acides aminés et notamment une liaison de type peptidique, les n unités (P-K) étant liées entre elles également par de telles liaisons par exemple de type peptidique.

Un oligonucléotide selon l'invention est donc caractérisé selon un premier mode de réalisation, en ce qu'il comprend une séquence codant pour une suite de motifs répétés comportant deux acides aminés.

Les codons de l'oligonucléotide peuvent être identiques pour tous les résidus Proline et/ou pour tous les résidus lysine. Ils peuvent aussi être différents pour un même résidu d'acide aminé, la variation tenant compte de la dégénérescence du code génétique.

De préférence, cet oligonucléotide est formé par une séquence codant pour plus de 2 unités (P-K). De préférence n est inférieur ou égal à 30, en particulier inférieur à 20 et avantageusement n est égal à 4, 5, 6, 7, 8, 9 ou 10 ou 15.

Les "oligonucléotides" selon l'invention peuvent être synthétisés chimiquement par toute technique disponible.

Le terme "polypeptide" par référence à l'enchaînement  $(P-K)_n$  désigne, dans le cadre de la présente demande, une séquence d'acides aminés comportant plus de 2 résidus d'acides aminés et pouvant comporter jusqu'à 60 résidus d'acides aminés.

Selon une première variante de réalisation de l'invention, l'oligonucléotide comprend plusieurs enchaînements codant pour un polypeptide répondant à la formule  $(P-K)_n$ , identiques ou différents, associés en tandem.

Ces oligonucléotides sont soit des répétitions d'un même enchaînement, soit des associations d'enchaînement différents. Le nombre d'enchaînements associés est variable, pouvant par exemple être compris entre 2 et 10 enchaînements.

5 Selon une autre variante de réalisation de l'invention, l'oligonucléotide précédemment défini est caractérisé en ce qu'il comprend au moins un enchaînement codant pour un polypeptide répondant à la formule  $(P-K)_n$ , dans laquelle la séquence des  $n$  unités  $(P-K)$  est interrompue par un ou plusieurs résidus d'acides aminés différents des  
10 résidus P ou K.

De préférence, les acides aminés supplémentaires incorporés dans la séquence formée par les unités  $(P-K)$  sont choisis pour ne pas modifier l'organisation du polypeptide codé par l'oligonucléotide, ou à tout le moins pour ne pas provoquer d'interaction avec des acides aminés d'une protéine  
15 dans laquelle ledit polypeptide serait incorporé, dans des conditions qui affecteraient la structure et/ou la fonction et/ou la localisation de cette protéine.

Ceci peut en particulier être le cas lorsque le nombre d'unités  $(P-K)$  est important ou que plusieurs enchaînements formés de séquences  
20 codant pour des motifs  $(P-K)_n$  sont associés en tandem et que la préparation de l'oligonucléotide correspondant nécessite que soient synthétisées plusieurs séquences nucléotidiques qui sont ensuite associées par exemple au moyen de séquences de liaison (linkers).

Selon un autre mode de réalisation de l'invention, l'oligonucléotide  
25 est tel que l'enchaînement codant pour le polypeptide comprenant les  $n$  unités  $(P-K)$  est complété à son extrémité 5' et/ou à son extrémité 3', par un ou plusieurs codons, codant par exemple pour au moins un résidu lysine à l'extrémité N-terminale du polypeptide.

A titre d'exemple, un oligonucléotide préféré selon l'invention est  
30 caractérisé en ce qu'il code pour un polypeptide répondant à la formule  $(P-K)$ , à la formule  $K-(P-K)_4$ , ou à la formule  $2K(P-K)_4$ .

Selon un mode de réalisation particulier, la composition de cet oligonucléotide correspond à l'une des séquences décrites dans les pages qui suivent et identifiées par les désignations SEQ ID No:1 et SEQ ID No:2.

5 Les oligonucléotides précédemment décrits constituent des moyens de base pour réaliser des séquences de nucléotides recombinantes capables d'exprimer des protéines ou des polypeptides de réserve de plantes enrichis en lysine.

L'invention a donc aussi pour objet une séquence de nucléotides recombinante comprenant un enchaînement de nucléotides codant pour  
10 une protéine de réserve d'une plante, caractérisée en ce qu'elle comprend en outre un oligonucléotide selon l'invention, inséré en un site de l'enchaînement de nucléotides choisi de façon telle que :

- l'expression de la séquence de nucléotides dans une cellule  
15 végétale déterminée permet d'obtenir une protéine de réserve modifiée localisée de façon identique ou similaire dans cette cellule à la protéine de réserve normale qui serait exprimée dans la même cellule, dans les mêmes conditions par l'enchaînement de nucléotides normal codant correspondant, et/ou

20 la protéine de réserve modifiée codée par la séquence de nucléotides est reconnue immunologiquement par des anticorps produits contre la protéine de réserve normale correspondante.

En particulier, les anticorps précités sont constitués par un sérum polyclonal ou sont obtenus contre des épitopes de la protéine de réserve  
25 normale qui sont conservés dans la protéine de réserve modifiée.

Les cellules végétales auxquelles il est fait référence ci-dessus comprennent toute cellule végétale, quelle que soit son origine tissulaire ou sa nature. En particulier on s'intéresse dans le cadre de l'invention aux  
30 cellules des organes de réserve, mais aussi aux cellules de feuilles, tiges, tubercules ...

Par l'expression "protéine de réserve" d'une plante, on entend dans le cadre de la présente demande une protéine synthétisée durant la maturation des semences et qui est utilisée pendant la phase de germination comme réserve principale de nutrition.

5 De façon générale, il s'agit d'un polypeptide susceptible d'être synthétisé dans un tissu de réserve quelle qu'en soit la localisation dans la plante ; les protéines de réserve mises en oeuvre dans le cadre de la présente demande sont en particulier celles produites dans les graines ou les semences des plantes de la famille des céréales, des crucifères ou des  
10 légumineuses, et sont par exemple les prolamines, les zéines.

Le choix du/des site(s) d'insertion de l'oligonucléotide dans l'enchaînement codant pour la protéine de réserve de plante est déterminé par le respect des conditions ci-dessus énoncées. Selon le cas, l'insertion peut avoir lieu dans un domaine répétitif (en termes de séquence d'acides  
15 aminés) de la protéine ou au niveau d'une extrémité C- ou N-terminale.

La condition donnée ci-dessus selon laquelle l'expression de la séquence de nucléotides recombinante de l'invention dans une cellule végétale permet d'obtenir une protéine de réserve modifiée, localisée  
20 identiquement ou similairement à la protéine de réserve normale qui serait exprimée dans les mêmes conditions dans la même cellule végétale, comporte par exemple pour les  $\gamma$ -zéines synthétisées, la possibilité d'être accumulées dans le réticulum endoplasmique des cellules végétales qui l'expriment, et en particulier dans les corps protéiques formés à partir de ce réticulum endoplasmique, lorsque la protéine est exprimée dans les  
25 cellules de l'endosperme.

Pour obtenir ce résultat au moyen des séquences de nucléotides recombinantes de l'invention, des systèmes d'expression adaptés à l'hôte cellulaire dans lequel est exprimée la séquence de nucléotides  
30 sélectionnée, et notamment les éléments de régulation, par exemple les promoteurs, sont choisis pour leur caractère fonctionnel dans le tissu contenant les cellules transformées. Des tests pour effectuer ce choix

peuvent être réalisés en s'appuyant sur les différentes constructions décrites dans les exemples.

5 Pour vérifier que les propriétés immunologiques de la protéine de réserve modifiée exprimée par la séquence de nucléotides de l'invention n'ont pas été modifiées de façon conséquente, on a utilisé par exemple des antisérums tels que l'antisérum  $\alpha$ G2, décrit précisément dans la partie expérimentale qui suit.

10 Selon un premier mode de réalisation de l'invention, la séquence de nucléotides recombinante est caractérisée en ce qu'elle est obtenue à partir d'un enchaînement codant de nucléotides qui conduit à l'expression d'une protéine de réserve naturellement pauvre en lysine.

15 De façon générale, cette séquence de nucléotides recombinante code pour une protéine de réserve modifiée dérivée d'une protéine de réserve naturellement produite par une plante utilisable en alimentation animale ou humaine.

20 Ainsi, les protéines de réserve dont on modifie la teneur en lysine dans le cadre de la présente invention sont avantageusement des protéines de réserve de plantes de la famille des céréales, des légumineuses ou des crucifères. Des protéines de réserve particulièrement importantes sont celles du maïs, en particulier les zéïnes, et plus spécialement la  $\gamma$ -zéïne du maïs, dont on cherche à augmenter la teneur en lysine.

25 Une séquence de nucléotides recombinante particulière selon l'invention est caractérisée en ce que l'enchaînement codant de nucléotides codant pour la  $\gamma$ -zéïne de maïs qu'elle contient, répond à la séquence présentée sur la Figure 9.

D'autres séquences de nucléotides recombinantes de l'invention sont caractérisées en ce que l'enchaînement codant de nucléotides qu'elles comprennent, code pour une protéine de réserve d'une plante

choisie parmi les suivantes : le soja, le tournesol, le tabac, le blé, l'avoine, la luzerne, le riz, le colza, le sorgo, Arabidopsis.

5 Selon un mode de réalisation préféré de l'invention, dans la séquence de nucléotides recombinante comprenant un enchaînement codant pour la  $\gamma$ -zéine de maïs, l'oligonucléotide de l'invention est inséré à la place de l'enchaînement codant pour le domaine Pro-X naturellement présent dans la séquence d'acides aminés de la  $\gamma$ -zéine du maïs ou à la suite de cet enchaînement. Le domaine Pro-X de la séquence d'acides aminés de la  $\gamma$ -zéine du maïs est constitué par les acides aminés situés  
10 entre les positions 70 et 91 de la séquence d'acides aminés représentée à la Figure 9, correspondant aux nucléotides 265 à 330 de la séquence représentée à la Figure 9.

De façon préférée, dans la séquence de nucléotides de l'invention, l'oligonucléotide à la place ou à la suite du domaine Pro-X présent entre  
15 les nucléotides 276 et 357 de la séquence représentée à la Figure 9.

Selon un autre mode de réalisation de l'invention, dans la séquence de nucléotides recombinante comprenant un enchaînement codant pour la  $\gamma$ -zéine de maïs, l'oligonucléotide de l'invention est inséré à la suite du domaine Pro-X conservé dans la séquence de la  $\gamma$ -zéine du maïs.

20 Selon une autre variante de réalisation, dans la séquence de nucléotides recombinante comprenant un enchaînement codant pour la  $\gamma$ -zéine de maïs, l'oligonucléotide de l'invention est inséré dans le domaine Pro-X maintenu dans la séquence de la  $\gamma$ -zéine.

Les insertions dont il est question ci-dessus sont réalisables par les techniques disponibles, par exemple par recombinaison des séquences ayant subi une ou plusieurs étapes de digestion enzymatique.

Dans un mode de réalisation particulier de l'invention, une protéine de réserve choisie enrichie en un acide aminé déterminé est exprimée dans des cellules végétales hétérologues. En d'autres termes une protéine  
30 de réserve présente naturellement dans une plante donnée, est exprimée

sous forme enrichie en acides aminés, dans une autre plante, ou dans une autre cellule que celle qui la produit naturellement.

Outre l'enchaînement codant pour une protéine de réserve de plante et l'oligonucléotide de l'invention, les séquences de nucléotides recombinautes de l'invention peuvent également comprendre un promoteur d'expression et, par exemple, un promoteur choisi pour son caractère spécifique pour l'expression dans certaines parties ou dans certains tissus des plantes, ou au contraire un promoteur choisi pour son caractère constitutif. A titre d'exemple, lorsqu'ils sont spécifiques, les promoteurs peuvent être spécifiques des semences et/ou d'organes ou de tissus déterminés de plantes. Ils peuvent alternativement ou également être spécifiques d'une phase de croissance, par exemple d'un stade déterminé de la germination.

A l'inverse l'utilisation de promoteurs constitutifs permet l'expression constante et générale de la protéine de réserve, entraînant une compétition entre l'expression de la protéine de réserve native lorsqu'elle est présente et la protéine de réserve modifiée.

A titre d'exemple, des promoteurs avantageux pour la réalisation de l'invention sont le promoteur de la  $\gamma$ -zéine du maïs contenu dans la séquence de 1,7 kb se trouvant en amont de la séquence codante représentée à la Figure 7, le promoteur du virus de la mosaïque du chou-fleur, à savoir le promoteur CaMV35S (EP-B-0131623), le promoteur constitutif du gène d'actine 1 de riz (PCT/US 9100073) ou le promoteur semence spécifique "High Molecular weight glutenine" du blé (Colot V. et al, 1987, EMBO Journal, vol. 6, p. 3559-3564).

Le cas échéant, ces promoteurs sont complétés par d'autres séquences de régulation, et en particulier des activateurs d'expression.

A titre d'exemple, d'autres promoteurs utilisables pour la réalisation de l'invention sont par exemple le promoteur du gène codant pour la protéine de réserve 2S des semences d'Arabidopsis thaliana, ou encore les promoteurs de la lectine de haricot ou de la  $\beta$ -phaséoline de haricot.

L'introduction supplémentaire d'activateurs de l'expression dans les séquences de régulation des séquences de nucléotides selon l'invention permet également d'augmenter le niveau de la transcription primaire de la séquence de nucléotides et, le cas échéant, d'augmenter la quantité de protéines de réserves modifiées produites. Des activateurs sont par exemple des introns de monocotylédones tels que l'intron 1 du gène d'actine de riz.

L'invention a également pour objet un vecteur de clonage et/ou d'expression, caractérisé en ce qu'il comprend, en un site non essentiel pour sa réplication, une séquence de nucléotides répondant à l'une des définitions précédemment données. Des vecteurs particuliers intéressants dans le cadre de la réalisation de l'invention sont par exemple les plasmides pP20 $\gamma$ Z, pH30 $\gamma$ Z ou pH45 $\gamma$ Z. Le plasmide pP20 $\gamma$ Z a été déposé à la CNCM (Paris, FRANCE) le 31 Octobre 1995 sous le numéro I-1640. Le plasmide pH45 $\gamma$ Z a été déposé à la CNCM le 31 Octobre 1995 sous le numéro I-1639.

Entre également dans le cadre de l'invention un polypeptide tel qu'exprimé par une séquence de nucléotides recombinante répondant aux définitions précédentes.

L'expression "polypeptide" n'introduit pas dans le cadre de la présente invention de limitation particulière quant au nombre d'acides aminés formant le polypeptide. Il peut donc s'agir de séquences comportant quelques acides aminés, habituellement désignées par l'expression "peptides", ou encore de séquences beaucoup plus longues telles que celles des protéines.

L'invention concerne à cet égard la  $\gamma$ -zéine de maïs modifiée riche en lysine, caractérisée en ce qu'elle est codée par une séquence de nucléotides recombinante décrite précédemment.

Selon un mode de réalisation préféré de l'invention, la  $\gamma$ -zéine de maïs modifiée enrichie en lysine est caractérisée en ce que sa séquence

d'acides aminés est modifiée par au moins un polypeptide répondant à la formule  $(P-K)_n$ , dans laquelle :

- n est un entier supérieur ou égal à 2,
- P représente un résidu d'acide aminé Proline,
- 5        - K représente un résidu d'acide aminé Lysine,
- le signe "-" symbolise une liaison entre les deux résidus d'acides aminés et notamment une liaison de type peptidique, les n unités (P-K) étant liées par des liaisons notamment de type peptidique.

10        Selon une variante de réalisation de l'invention le polypeptide intégré dans la séquence d'acides aminés de la  $\gamma$ -zéine répond à la formule  $K-(P-K)_n$ .

15        Les polypeptides de l'invention répondant à l'une des formules  $(P-K)_n$ ,  $K-(P-K)_n$  ou à des variantes sont substitués à une séquence naturellement présente dans la  $\gamma$ -zéine de maïs normale ou insérés avec délétion d'un ou plusieurs acides aminés de la séquence d'acides aminés de la  $\gamma$ -zéine de maïs normale, ou ajoutés dans la séquence d'acides aminés de la  $\gamma$ -zéine normale, le site d'insertion du polypeptide étant choisi de façon telle que :

20        - lorsque la  $\gamma$ -zéine modifiée riche en lysine est produite dans une cellule hôte, notamment une cellule végétale, elle est localisée de façon identique ou similaire dans cette cellule à la  $\gamma$ -zéine de maïs normale qui serait produite dans les mêmes conditions, dans la même cellule hôte et/ou,

25        - la  $\gamma$ -zéine de maïs modifiée est reconnue par des anticorps dirigés contre la  $\gamma$ -zéine normale de maïs.

Les protéines P20 $\gamma$ Z représentée à la figure 11 ou H30 $\gamma$ Z ou H45 $\gamma$ Z représentée à la figure 10, sont des exemples préférés de la réalisation de l'invention et représentent des  $\gamma$ -zéines de maïs modifiées enrichies en lysine.

L'invention a également pour objet une cellule hôte recombinante, caractérisée en ce qu'elle comprend une séquence de nucléotides décrite ci-dessus.

Des cellules hôtes intéressantes sont par exemple les cellules de bactéries, telles que celles de E. coli ou d'Agrobacterium tumefaciens. De préférence, dans le cadre de l'invention, et pour l'expression stable de la protéine de réserve modifiée recherchée, on aura recours à des cellules hôtes d'origine végétale.

A titre d'exemple, ces cellules d'origine végétale sont des cellules de semences, de plante, et par exemple, à titre préféré, des cellules d'endosperme de semences de maïs.

La séquence de nucléotides de l'invention est de préférence introduite dans le génome de la cellule hôte de façon stable et dans des conditions telles que la protéine de réserve exprimée enrichie en acides aminés et en particulier en lysine est localisée comme le serait la protéine correspondante normale dans la même cellule hôte.

Des techniques variées sont disponibles pour transformer les cellules hôtes. A titre d'exemple, et en vue de transformer de façon stable ou de façon transitoire les cellules hôtes, on utilisera les techniques d'électroporation, de bombardement de microprojectiles portant de l'ADN, via canon à particules, via culture d'explants avec Agrobacterium tumefaciens, par pénétration de microfibres.

Outre les cellules d'endosperme de semences de maïs, des cellules de soja, de tournesol, de tabac, de blé, d'avoine, de luzerne, de riz, de colza, de sorgo ou d'Arabidopsis peuvent être mises en oeuvre pour exprimer les séquences de nucléotides de l'invention.

La présente demande concerne aussi les semences produisant un polypeptide tel que décrit ci-dessus et les plantes produisant ce même polypeptide. De préférence, ces plantes sont du maïs.

L'invention a aussi pour objet des semences obtenues à partir des plantes transformées exprimant le polypeptide de l'invention, en d'autres

termes le polypeptide de réserve modifié enrichi en acides aminés déterminés.

Selon un mode de réalisation particulièrement intéressant de l'invention, les protéines  $\gamma$ -zéïnes modifiées enrichies en lysine, sont exprimées chez des mutants opaque 2 de maïs. Ces mutants o2 décrits par Emerson R.A. et al. (1935, Cornell Univ. Agric. Exp. Stn. Mem. 180) et caractérisés par Mertz E.T. et al (1964, Science 145 : 279-280) voient leur contenu en lysine fortement **augmenté** améliorant ainsi fortement les qualités nutritives du maïs (compensation de son bas niveau en cet acide aminé essentiel). Les maïs classiques ont un contenu en lysine d'environ 0,24% du produit brut (poids du grain total), les maïs opaque 2 avoisinent en revanche 0,5% en lysine. Cependant ils présentent des caractéristiques agronomiques insuffisantes car leur endosperme est beaucoup moins vitreux et se révèle très friable (phénotype "starchy"). Ceci a pour effet de les rendre extrêmement sensibles aux organismes pathogènes et aux étapes de traitements post recolte. Ce phénotype est en fait dû à la diminution importante de certaines protéines de réserve en particulier les alpha zéïnes. En fait Opaque 2 code pour un facteur de transcription nécessaire à l'expression de certains gènes zéïnes (Schmidt R. J. et al, 1990, Proc. Natl. Acad. Sci. USA 87, 46-50).

Des dérivés opaque 2 n'ayant plus les inconvénients précités ont été développés par amélioration génétique classique, il s'agit des maïs QPM (Quality Protein Maize). L'analyse génétique récente de ces maïs (Lopes M.A. et al, 1995, Theor. Appl. Genet. 19, 274-281) a montré que seulement 2 ou 3 loci se révèlent clefs dans ces modifications favorables. Des analyses génétiques et biochimiques plus poussées permettent de postuler qu'un des 3 loci responsables est le locus  $\gamma$ -zéïne : les génotypes de maïs qui portent une duplication de ce gène situé dans la région centrométrique du chromosome 7, s'avèrent tous être des modificateurs d'opaque 2 (Lopes M.A. et al, 1995, Mol. Gen. Genet. 19 : 247 : 603-613).

La présente invention permet de préparer des maïs mutants opaque 2 à partir de maïs n'ayant qu'un gène de  $\gamma$ -zéine au niveau du chromosome 7, qui sont complémentés par l'addition d'une séquence recombinante codant pour une  $\gamma$ -zéine de maïs enrichie en lysine. Outre le fait d'acquérir les propriétés de dureté similaire à un maïs opaque 2 non mutant, il a l'avantage d'augmenter significativement la teneur en lysine, dépassant ainsi celle des maïs QPM.

La présente invention permet d'obtenir des mutants opaque 2 de maïs modifiés, dans lesquels on a inséré une séquence de nucléotides recombinante codant pour une  $\gamma$ -zéine de maïs enrichie en lysine.

L'invention a aussi pour objet un procédé pour l'obtention de plantes ou de semences exprimant une protéine de réserve modifiée, caractérisé en ce qu'il comprend les étapes de :

a) transformation d'une cellule végétale, avec une séquence de nucléotides ou un vecteur ci-dessus décrits, dans des conditions permettant l'expression de façon stable et fonctionnelle de la protéine de réserve modifiée codée par la séquence de nucléotides;

b) régénération de plantes à partir de la cellule de plante transformée de l'étape a), pour obtenir des plantes exprimant la protéine de réserve modifiée,

c) le cas échéant, obtention des semences à partir des plantes modifiées obtenues à l'étape b).

Selon un mode de réalisation avantageux de l'invention, la plante transformée est le maïs et la protéine de réserve modifiée enrichie est la  $\gamma$ -zéine, enrichie en lysine.

L'invention concerne aussi les plantes obtenues par mise en oeuvre d'un tel procédé.

D'autres caractéristiques et avantages de l'invention apparaissent dans les exemples et dans les figures qui suivent.

**Figure 1**

- Carte de restriction du plasmide pP20 $\gamma$ Z

**Figure 2**

5 - Carte de restriction du plasmide pP45 $\gamma$ Z

**Figure 3**

10 - Représentation schématique des protéines codées par les gènes modifiés et non modifiés de la  $\gamma$ -zéine :  $\gamma$ -zéine de type sauvage ( $\gamma$ Z),  $\gamma$ -zéines riches en lysine (P20 $\gamma$ Z, H30 $\gamma$ Z, H45 $\gamma$ Z et N13 $\gamma$ Z) résultant de l'insertion des oligonucléotides codant pour des séquences riches en lysine. La séquence d'acides aminés des polypeptides insérés est indiquée en utilisant le code à une lettre de la représentation des acides aminés. Les abréviations suivantes sont utilisées :

15 Term : terminal ;

ProX DOMAIN : domaine linker proline-Xaa.

**Figure 4** - Analyse in-vitro des  $\gamma$ -zéines riches en lysine. (A) traduction in-vitro et translocation des transcrits correspondant aux  $\gamma$ -zéines modifiées riches en lysine ; lignes 1, 5, 9 et 13 : produits complets de traduction ; lignes 2, 6, 10 et 14 : produits complets de traduction après translocation dans les microsomes canins (CM) ; lignes 3, 7, 11 et 15 : produits de translocation résistant à l'action de la protéinase K (PK) ; lignes 4, 8, 12 et 16 : totalité des produits de traduction après traitement avec la protéinase K en présence de 0,5% Nonidet P40(NP40). (B) Immunoprécipitation des produits de traduction in-vitro correspondant aux protéines  $\gamma$ -zéine et  $\gamma$ -zéine modifiées riches en lysine, en utilisant l'anti-sérum  $\alpha$ PL. Ligne 1 :  $\gamma$ -zéine ; ligne 2 : P20 $\gamma$ Z ; ligne 3 : H30 $\gamma$ Z ; ligne 4 : H45 $\gamma$ Z et ligne 5 : N13 $\gamma$ Z. (C) Même légende que pour (B) mais en utilisant l'antisérum  $\alpha$ G2. Les

marqueurs de poids moléculaire (en kilodaltons) sont indiqués sur la gauche.

**Figure 5** - Activité tissu-spécifique du promoteur de la  $\gamma$ -zéine. Des endospermes de maïs, des embryons et des feuilles de maïs ont été co-transformés par bombardement de particules en utilisant les constructions représentées dans la figure (dans la partie à droite). Les activités relatives de la luciférase (LUC, colonnes en gris) et de la  $\beta$ -glucuronidase (GUS, colonnes hachurées) ont été exprimées sous la forme d'un multiplicateur des valeurs obtenues avec des projectiles nus  $\pm$  la déviation standard des différents ratios.

**Figure 6** - Expression des  $\gamma$ -zéines riches en lysine dans l'endosperme sub-aleuronique des cellules. (A) Immunoblot avec un antisérum  $\alpha$ PL, de protéines extraites d'endospermes transformés par pN13 $\gamma$ Z (ligne 2), pH45 $\gamma$ Z (ligne 3), pH30 $\gamma$ Z (ligne 4) et pP20 $\gamma$ Z (ligne 5). Le témoin (ligne 1) correspond à des endospermes non transformés. Les marqueurs de poids moléculaire (en kilodaltons) sont indiqués à gauche. (B) expression des transcrits H45 $\gamma$ Z et N13 $\gamma$ Z dans des endospermes transformés de façon transitoire. Les cDNAs obtenus à partir des tissus transformés avec pH45 $\gamma$ Z (ligne 2), pN13 $\gamma$ Z (ligne 3) et le témoin (ligne 1) ont été amplifiés par PCR et analysés en utilisant un oligonucléotide synthétique codant pour une séquence riche en lysine à titre de sonde.

**Figure 7** - Accumulation de  $\gamma$ -zéines riches en lysine dans les corps protéiques de l'endosperme. (A) Analyse en immunoblot en utilisant l'antisérum  $\alpha$ PL, des corps protéiques isolés à partir des endospermes transformés avec pP20 $\gamma$ Z (ligne 1), pH30 $\gamma$ Z (ligne 2), pH45 $\gamma$ Z (ligne 3), pN13 $\gamma$ Z (ligne 4) et aucun ADN (ligne 5). (B) Analyse en immunoblot en utilisant l'antisérum  $\alpha$ PL, des corps protéiques isolés à partir des

endospermes transformés avec pP20 $\gamma$ Z, pH30 $\gamma$ Z et pH45 $\gamma$ Z et digérés avec la protéinase K en présence d'un tampon isotonique (Sucr., lignes 1, 3 et 5) ou un tampon hypotonique (H<sub>2</sub>O, lignes 2, 4 et 6). Les marqueurs de poids moléculaire (en kilodaltons) sont indiqués sur la gauche.

5

**Figure 8** - Co-localisation des protéines P20 $\gamma$ Z avec les  $\alpha$ - et les  $\gamma$ -zéines dans les corps protéiques de l'endosperme du maïs. Une analyse en immunocytochimie a été réalisée sur des sections ultrafines en utilisant des anticorps  $\alpha$ PL (marqués avec des particules d'or de diamètre 15 nm) et des anticorps  $\alpha$ Z et  $\alpha$ G2 (marqués avec des particules d'or de 5 nm). (A) Corps protéiques de l'endosperme transformé avec pP20 $\gamma$ Z, immunomarqués avec l'anticorps  $\alpha$ PL. (B) Immunolocalisation de P20 $\gamma$ Z (marqué avec des particules d'or de 15 nm) et de l' $\alpha$ -zéine (marquée avec les particules d'or de 5 nm) dans les corps protéiques isolés à partir des endospermes transformés avec pP20 $\gamma$ Z. (C) et (D) Immunolocalisation de P20 $\gamma$ Z (marqué avec les particules d'or de 15 nm) et des  $\gamma$ -zéines (marquées avec les particules d'or de 5 nm) dans des corps protéiques isolés à partir des endospermes transformés avec pP20 $\gamma$ Z. Les flèches indiquent les sections tangentielles des corps protéiques.

20

**Figure 9** - Séquence codante de l'ADNc la  $\gamma$ -zéine de maïs et la séquence d'acides aminés correspondante.

25

**Figure 10** - Séquence codante de l'ADNc de la H45 $\gamma$ Z zéine de maïs et la séquence d'acides aminés correspondante.

La séquence riche en lysine (28 acides aminés) est introduite entre les résidus d'acides aminés 92 et 119 de la séquence représentée à la Figure 10.

30

**Figure 11** - Séquence codante de l'ADNc de la P20 $\gamma$ Z zéine de maïs et la séquence d'acides aminés correspondante.

La séquence riche en lysine (14 acides aminés) est introduite entre les résidus d'acides aminés 92 et 119 de la séquence représentée à la Figure 11.

**Figure 12** - Cartes de restriction des plasmides pBin 19P20 $\gamma$ Z et pBin 19H30 $\gamma$ Z.

### EXEMPLES

#### A) Préparation de $\gamma$ -zéines modifiées enrichies en lysine et expression de ces protéines modifiées suivie de l'accumulation dans les corps protéiques des cellules de l'endosperme de maïs

La  $\gamma$ -zéine est une protéine de réserve du maïs, ayant un poids moléculaire de 28 kD, riche en soufre, qui est accumulée dans les cellules de l'endosperme avec les  $\alpha$ -et  $\beta$ -zéines, dans les corps protéiques dérivés du réticulum endoplasmique granulaire (ER) (Ludevid et al., 1984, Plant Mol. Biol. 3, 227-234 ; Lending et al., 1984, Plant Cell 1, 1011-1023). La séquence d'acides aminés déduite de la séquence nucléotidique de l'ADNc (Prat et al., 1985, Nucl. Acids Res. 13, 1493-1504) et des clones génomiques (Boronat et al., 1986, Plant Sci. 47, 95-102) montre que la  $\gamma$ -zéine n'a pas d'homologie avec les polypeptides de type  $\alpha$ -zéine. Bien que la  $\gamma$ -zéine soit codée par 1 à 2 gènes par génome haploïde (Boronat et al., 1986, Plant Sci. 47, 95-102) elle représente 10-15% de la totalité des protéines des endospermes de maïs. L'expression du gène de la  $\gamma$ -zéine dans des systèmes hétérologues tels que les oocytes de *Xenopus* (Torrent et al., 1994, Planta 192, 512-518) et dans *Arabidopsis thaliana* (Geli et al., 1994, Plant Cell 6, 1911-1922), indique que les polypeptides de  $\gamma$ -zéine restent stables et sont capables en tant que tels de former des corps

protéiques dérivés du réticulum endoplasmique, à l'intérieur des cellules. De plus, des analyses par délétion de différents domaines structuraux de la  $\gamma$ -zéine ont montré que la séquence N-terminale incluant le domaine répété riche en proline, était responsable de la rétention de la  $\gamma$ -zéine dans le réticulum endoplasmique et que le domaine C-terminal riche en cystéine était responsable de la formation des corps protéiques. Le domaine Pro-X ne semblait pas en relation avec la stabilité de la protéine ni avec sa localisation ciblée (Geli et al., 1994, Plant Cell 6, 1911-1922).

### Matériel et Méthodes

#### **Matériel végétal**

Après une stérilisation en surface (1) des graines de stade 17 DAP ("days after pollination") du maïs de sélection W64A ont été disséquées à la main et la couche de péricarpe et d'aleurone a été séparée des endospermes. Des sections tangentiellles ont été réalisées de façon à exposer une grande partie de la surface sub-aleuronique. Si nécessaire, des embryons ont été isolés et des feuilles de plantes âgées de 7 jours ont été disséquées pour retirer le tissu épidermique. Après dissection, les échantillons ont été déposés dans des boîtes de Petri sur des papiers filtres humidifiés avec un milieu MS (Murashige et Skoog, 1962, Physiol. Plant. 15, 473-497).

#### **Constructions plasmidiques**

Un premier ensemble de plasmides, pKSG2, pHbP2, pPbP4 et pNaN1 a été obtenu pour permettre l'introduction de sites de restriction dans le gène codant pour la  $\gamma$ -zéine. pKSG2 et pHbP2 ont été construits selon la description décrite dans la publication de Torrent M. et al. (Planta

(1994) 192 : 512-518). Le plasmide pKSG2 contient la séquence codante de la  $\gamma$ -zéine.

Le plasmide pHbP2 est obtenu à partir de pKSG2 et contient une séquence codant pour une  $\gamma$ -zéine mutée dans laquelle le domaine Pro-X de la protéine a été délété.

Le plasmide pPbP4 a été obtenu à la suite de deux étapes de clonage : (i) le fragment de restriction Sall-PvuII de 350 pb de pKSG2 a été cloné dans un plasmide Bluescript (pBSKS, Stratagene, La Jolla, California USA) restreint avec Sall et EcoRV (pKSC4) et (ii) le fragment de restriction PvuII-XbaI de 600 bp de pKSG2 a été cloné dans les sites de restriction SmaI-XbaI de pKSC4. La nouvelle construction pPbP4 contient un site de restriction utile EcoRI juste avant le domaine P-X de la séquence codante de la  $\gamma$ -zéine.

Le plasmide pNaN1 a également été obtenu à l'aide de deux étapes de clonage : (i) le fragment NaeI-XbaI de 250 bp de pKSG2 a été cloné dans le plasmide pBSKS restreint avec EcoRV-XbaI (pKSC8) et (ii) le fragment de restriction NaeI-HindIII de 700 bp (à extrémités franches) de pKSG2 a été cloné dans le site de restriction HindIII de pKSC8. La nouvelle construction, pNaN1, contient des sites de restriction ClaI et HindIII à une position située à 15 nucléotides avant le codon stop de la  $\gamma$ -zéine.

Deux oligonucléotides synthétiques répondant aux séquences suivantes : SEQ ID N°1 :

5'CGATGAATTCAAACCAAAGCCAAAGCCGAAGCCAAAAGAATTCA3', et sa séquence inverse appelée SEQ ID N°2, dont la séquence est la suivante :

5'AGCTTGAATTCTTTTGGCTTCGGCTTTGGCTTTGGTTTGAATTCAT3'

codant pour des séquences riches en lysine désignées K(P-K)<sub>4</sub>, ont été hybridées, digérées avec EcoRI et clonées dans un site EcoRI de pHbP2 et pPbP4. Trois clones ont été sélectionnés : pPo2 et pHo3 contenant la

séquence codante pour  $K(P-K)_4$  et pHo4 comprenant la forme tronquée de la séquence codant pour le  $\gamma$ -zéine contenant un tandem  $2K(P-K)_4$  (sous forme d'une séquence  $K(P-K)_4$  EF  $K(P-K)_4$ ) de la séquence codante riche en lysine. Les mêmes oligonucléotides hybridés ont été digérés avec des enzymes ClaI-HindIII et clonés dans le plasmide pNaN1 restreint à l'aide des mêmes enzymes. Le clone sélectionné, pNo1, contenait la séquence codant pour la séquence riche en lysine  $K(P-K)_4$  à l'extrémité N-terminale de la  $\gamma$ -zéine modifiée correspondante.

Pour la transformation transitoire de l'endosperme, les séquences codant pour la  $\gamma$ -zéine modifiée de pPo2 et pHo3 ont été insérées sous la forme de fragments HincII-NheI dans des sites SmaI-XbaI de pDH51 (Pietrzah et al., 1986, Nucl. Acid Res. 14, 5857-5868) contenant le promoteur 35S du virus de la Mosaïque du chou-fleur (CaMV). Le promoteur pP20 $\gamma$ Z obtenu par l'insertion décrite ci-dessus des fragments HincII-NheI dans les plasmide pDH51 contient la séquence codante de la  $\gamma$ -zéine enrichie en lysine (Figure 8) ainsi que les signaux de la séquence 35S du virus CaMV pour la formation de l'extrémité 3' et la polyadénylation. La séquence codante chimère P20 $\gamma$ Z a été construite à partir de la région codante de la  $\gamma$ -zéine contenu dans le plasmide pKSG2 après différentes étapes de clonage. Le promoteur de 1,7 kb de la  $\gamma$ -zéine (Reina et al. 1990, Nucl. Acids Res. 18, 6426) a été inséré dans les extrémités franches d'un fragment HindII-PvuI dans pHo4 et pNo1 restreints avec XhoI et obtenus avec des extrémités franches. Les constructions pH45 $\gamma$ Z et pN13 $\gamma$ Z ont respectivement été obtenues.

Les nouvelles constructions, respectivement appelées pP20 $\gamma$ -Z, pH30 $\gamma$ Z, pH45 $\gamma$ Z et pN13 $\gamma$ Z ont été utilisées dans des expériences de bombardement biolistique.

Pour étudier la spécificité de différents promoteurs vis-à-vis des tissus végétaux, deux constructions, p1.7 $\gamma$ ZGUS et pCaMV35SLUC ont été utilisées. p1.7 $\gamma$ ZGUS a été obtenue par insertion du promoteur de la  $\gamma$ -

zéine de 1,7 kb (HindIII-PvuI) dans un plasmide dérivé de pPuC18 contenant le gène GUS et les signaux NOS de polyadénilation en 3' de pBI 101.1 (Jefferson et al., 1987, Embo. J. 6, 3901-3907). pCaMV35SLUC a été obtenu par insertion du gène codant pour la luciférase (LUC) de pAHC18 (Bruce et al., 1989, P. H. 86, 9692-9696) dans le polylinker de pDH51 (Pietrzak et al., 1986, Nucl. Acids Res. 18, 6426).

### **Analyse in-vitro**

Les plasmides dérivés de pBSKS contenant les séquences codantes de la  $\gamma$ -zéine (pKSG2) et de la  $\gamma$ -zéine riche en lysine (pPo2, pHo3, pHo4 et pNo1) ont été transcrits in-vitro suivant les protocoles standards (Sambrook et al., 1989, Molecular Cloning: A laboratory manual, 2nd. ed., Cold Spring Harbor Laboratory Ed., Cold Spring Harbor, New York). Une traduction in-vitro et une translocation des transcrits synthétiques ont été réalisées selon la technique de Torrent et al., (1994, Panta 192, 512-518), à l'exception des microsomes canins (CM) utilisés qui provenaient de Promega (Madison, Wis., USA). L'immunoprécipitation des produits traduits a été effectuée essentiellement selon la méthode de Borgese et Gaetani (1980) en utilisant un sérum de lapin anti- $\gamma$ -zéine $\alpha$ -G2, (Ludevid et al., 1985, Plant Sci. 41, 41-48) et un antisérum  $\alpha$ PL.  $\alpha$ PL est un antisérum polyclonal de lapin obtenu contre le peptide synthétique EFK(P-K)<sub>8</sub>EF. Ce peptide a été synthétisé en utilisant la technique de synthèse en phase solide décrite par Celma et al., 1992.

### **Bombardement de microprojectile**

L'ADN plasmidique a été absorbé sur des particules d'or (1,0  $\mu$ m, Bio-Rad, Lab., Richmond, CA, USA) selon un protocole décrit par Kikkert (Plant Cell, 33:221-226, 1993). Toutes les cibles ont été bombardées deux

fois, en utilisant le dispositif BioRad Biolistic PDS-1000/He. Les cibles ont été positionnées 8 cm derrière un écran stoppant les macroporteurs, qui étaient positionnés 1 cm au dessous du disque de rupture 900 PSI. Après le bombardement, les échantillons ont été incubés pendant 24 heures à 26°C dans le noir. Des témoins étaient constitués par les cibles bombardées avec des microprojectiles dépourvus d'ADN.

### Tests enzymatiques

Les tissus bombardés avec les plasmides p1.7 $\gamma$ ZGUS et pCaMV35SLUC ont été homogénéisés sur de la glace dans un tampon contenant 25mM Tris, à pH 7,8, 2mM DTT, 10% de glycérol et 1% de Triton X-100. Après centrifugation à 12.000xg pendant 5 minutes, les surnageants ont été décantés et la protéine soluble totale dans les extraits a été quantifiée en utilisant le test Bradford (Bio-Rad). L'activité GUS a été testée par analyse fluorimétrique selon la description de Jefferson (1987) en utilisant la 4-méthyl-ombélliféryl- $\beta$ -D-glucuronide (MUG) en tant que substrat. L'activité LUC a été déterminée en utilisant le système de test Luciférase (Luciferase Assay System Kit) diffusé par Promega, selon les instructions du fabricant.

### Extraction des protéines de réserve et analyse sur gel des protéines

Des endospermes transformés avec pP20 $\gamma$ Z, pH20 $\gamma$ Z, pH45 $\gamma$ Z et pN13 $\gamma$ Z ont été réduits à l'état de farine et les  $\alpha$ -zéïnes ont été extraites à l'aide de trois séries de solvants contenant 70% d'éthanol. La farine résiduelle a été séchée à l'air, et les protéines totales ont été extraites avec un tampon contenant 0,25 M Tris-HCl, pH 6,8, 4% sodium dodecyl-sulfate (SDS) et 5% 2-mercaptoethanol, pendant 1 heure à température ambiante. Les extraits de protéines ont été analysés par SDS-PAGE et

immunoblot selon la description de Ludevid et al., 1985. Des feuilles de nitrocellulose ont été intubées avec de l'antisérum  $\alpha$ PL (dilution 1:500) et de la peroxidase de Raifort conjuguée avec un anticorps secondaire (ECL Western Blotting System, Amersham, Buckinghamshire, UK) a été utilisée pour la détection de la protéine.

### **Analyse de l'expression de l'ARN**

L'ARN total a été extrait selon la description de Logemas et al., 1987. L'ADN complémentaire (ADNc) a été préparé en utilisant la transcriptase reverse et l'oligo dT de Bibco BRL (Gaithersburg, MD, USA) selon les instructions du fabricant et cet ARN a été amplifié par une réaction de PCR. Des oligonucléotides amorces utilisées pour la PCR étaient des séquences de 20-mers correspondant aux extrémités de 5' et 3' du gène de structure de la  $\gamma$ -zéine. Des protocoles standards ont été utilisés pour la préparation de sondes marquées au  $^{32}$ P, et pour l'analyse sur gel de l'ADN (Sambrook et al. Molecular Cloning: A laboratory manual, 2nd. ed., Cold Spring Harbor Laboratory Ed., Cold Spring Harbor, New York) en utilisant un oligonucléotide synthétique codant pour une séquence riche en lysine (voir ci-dessus) en tant que sonde.

### **Isolement des corps protéiques et traitement par la protéase**

Ces protocoles ont été décrits antérieurement (Torrent et al., Planta, 180: 90-95, 1989).

### **Microscopie électronique**

Les corps protéiques des endospermes de type sauvage et des endospermes transformés par pP20 $\gamma$ Z ont été fixés avec 2,5% de

paraformaldéhyde dans 20 mM de tampon phosphate à pH7,2, pendant 1 heure à température ambiante, et transformés selon la description de Geli et al., 1994, Plant Cell 6, 1911-1922) toutefois en utilisant un antisérum  $\alpha$ -PL et un colloïde d'or et de protéine A ayant un diamètre de 15 nm. Pour le marquage double, des sections ultra fines ont été d'abord incubées avec  $\alpha$ -PL et le colloïde d'or et de protéine A (15 nm de diamètre) a été utilisé pour la détection de l'anticorps. Après lavage, des sections ont été incubées avec 0,15 mg/mL de protéines A pendant 20 minutes pour saturer les immunoglobulines et finalement les grilles ont été incubées avec des séra  $\alpha$ -G2 ou  $\alpha$ -Z1 et le colloïde or/protéine A (diamètre 5 nm) a été utilisé pour la détection de l'anticorps.  $\alpha$ -Z1 est un antisérum polyclonal de lapin dirigé contre l' $\alpha$ -zéine obtenu selon la description de Ludevid et al., 1985, Plant Sci. 41, 41-48).

## Résultats

### Construction des $\gamma$ -zéines riches en lysine

Les inventeurs avaient démontré l'importance du domaine répété riche en proline ("proline-rich repeat") et du domaine C-terminal riche en cystéine pour la rétention dans le réticulum endoplasmique de la  $\gamma$ -zéine et la formation des corps protéiques contenant ces protéines dans les cellules de feuilles de *Arabidopsis* (Geli et al., 1994, Plant Cell 6, 1911-1922). Sur la base de ces résultats préalables, la possibilité d'insérer des séquences riches en lysine dans différents domaines de la  $\gamma$ -zéine, de façon à créer une  $\gamma$ -zéine modifiée correctement ciblée et accumulée dans les cellules de l'endosperme a été investiguée afin d'améliorer les qualités nutritives du maïs.

Les inventeurs ont à présent construit des gènes modifiés de la  $\gamma$ -zéine par introduction d'oligonucléotides synthétiques codant pour des

séquences riches en lysine dans différents sites de la séquence codante de la  $\gamma$ -zéine. Des constructions de  $\gamma$ -zéine modifiée ont été créées de façon à éviter le placement des séquences codantes riches en lysine dans les domaines constitués par le domaine répété en tandem et le domaine riche en cystéine. Des modifications de la séquence codante de la  $\gamma$ -zéine ont été effectuées dans la séquence correspondant au domaine Pro-X. De plus pour minimiser une altération du repliement de la protéine, les séquences riches en lysine  $(P-K)_n$  ont été définies pour imiter la séquence du domaine Pro-X. Comme on le voit sur la figure 3, dans la protéine P20 $\gamma$ Z, une séquence  $K(P-K)_4$  a été introduite après la région Pro-X et dans la protéine H30 $\gamma$ Z et dans la protéine H45 $\gamma$ Z, des séquences d'acides aminés comprenant  $K(P-K)_4$  et  $2K(P-K)_4$  respectivement remplacent le domaine Pro-X de la  $\gamma$ -zéine ( $\gamma$ Z, Fig. 3). Pour explorer si l'extrémité C-terminale était un site permissif pour l'introduction des séquences riches en lysine, une protéine supplémentaire N13 $\gamma$ Z, a été créée par l'insertion d'une séquence contenant  $K(P-K)_4$  cinq acides aminés en amont de l'extrémité C-terminale (Fig.3).

#### Activité du promoteur de la $\gamma$ -zéine dans l'endosperme de maïs transformé

Pour déterminer si des  $\gamma$ -zéines riches en lysine pouvaient être exprimées dans les cellules de l'endosperme, on a tout d'abord recherché un promoteur efficace et un système de transformation. Un promoteur de la  $\gamma$ -zéine, contenant une séquence en amont de 1,7 Kb (Reina et al., 1990, Nucleic Acid Research, vol. 18, p. 6426) et le promoteur CaMV contenant 625 bp de la séquence en amont de la protéine 35S du virus de la mosaïque du chou-fleur CaMV ont été testés. Jusqu'à présent aucune information n'était disponible sur l'analyse fonctionnelle des fusions de gènes avec le promoteur de la  $\gamma$ -zéine dans les plantes monocotylédones transgéniques. Pour analyser l'activité et la spécificité tissulaire du

promoteur de la  $\gamma$ -zéine, deux gènes chimères ont été construits (voir Fig. 5). l'expression transitoire par bombardement biolistique (Klein et al., 1988 PNAS 85:4305) a été utilisée comme procédure de transformation du maïs pour l'analyse du promoteur aussi bien que pour des expériences d'expression des  $\gamma$ -zéines riches en lysine. Des endospermes de maïs de stade 17 DAP (17 jours après la polynisation) (le péricarpe et les cellules de la couche aleuroniques étaient enlevés), des embryons (17 DAP) et des jeunes feuilles (âgées de 10 jours) ont été bombardés avec des projectiles d'or recouverts avec de l'ADN plasmidique contenant les deux constructions. La figure 5 montre les activités de  $\beta$ -glucuronidase (GUS) et de luciférase (LUC) présentes dans les trois tissus testés : l'endosperme, l'embryon et la feuille, par rapport à l'expérience témoin. On doit noter que les résultats correspondent à la moyenne d'au moins 3 expériences indépendantes réalisées. Toute l'activité GUS sous le contrôle du promoteur de la  $\gamma$ -zéine était restreinte à l'endosperme, puisque aucune expression GUS n'a été détectée dans l'embryon et dans les feuilles. De plus, les endospermes bombardés ont été colorés histochimiquement pour déterminer le nombre de clusters cellulaires exprimant la protéine GUS. Le profil de coloration corroborait les résultats précédents, GUS était fortement exprimé dans les endospermes (le nombre moyen de clusters cellulaires colorés pour GUS par endospermes était de 150) et des tâches bleues n'ont pas été détectées dans l'embryon et dans les feuilles. Au contraire, le promoteur CaMV 35S conférait une activité LUC dans tous les tissus testés (Fig. 5), mais il existait cependant des différences quantitatives entre l'activité relative de l'enzyme dans les feuilles et les embryons par rapport à l'endosperme. Ces différences pouvaient être attribuées à une variabilité intrinsèque de l'activité constitutive du promoteur CaMV parmi les différents tissus du maïs ou à une pénétration faible des particules recouvertes par l'ADN dans les cellules mésophyles contenant un système vacuolaire important. Il existe dans l'état de la

technique des preuves selon lesquelles le promoteur CaMV aurait habituellement une faible activité dans des cellules de plantes monocotylédones (Fromm et al., 1985, Proc. Natl. Acad. Sci. USA 82, 5824-5828) ; cependant les inventeurs ont montré ici une forte activité du promoteur CaMV dans des cellules d'endospermes. Ceci a permis de conclure que l'activité du promoteur de la  $\gamma$ -zéine et du promoteur CaMV 35S était très forte dans les endospermes de maïs et ainsi que les deux promoteurs pouvaient s'avérer utiles pour contrôler l'expression de la protéine dans ce tissu.

Pour déterminer si les protéines mutantes codées par les constructions étaient compétentes avec la fonction de translocation membranaire, des expériences de transcription-traduction in-vitro ont été effectuées en présence de microsomes pancréatiques de chien. Les transcrits synthétiques de chaque construction ont été traduits et la translocation à travers les membranes de microsomes testée par recherche de la protection vis-à-vis de la digestion par la protéinase K. Ces résultats rapportés à la figure 4A indiquent que les poids moléculaires apparents des polypeptides synthétisés in-vitro reflètent les mutations introduites (Fig. 4A, lignes 1, 5, 9 et 13). En présence des microsomes et de la protéinase K, des peptides de faible poids moléculaire n'ont pas été observés, indiquant que les chaînes polypeptidiques complètes des  $\gamma$ -zéines modifiées étaient transportées à travers les membranes microsomales (Fig. 4A, lignes 3, 7, 11 et 15). En comparant le résultat de la translocation des quatre  $\gamma$ -zéine modifiées, on a observé que la protéine H45 $\gamma$ Z (Fig. 4A, ligne 11), qui contenait l'insertion de 10 acides aminés de type lysine, a subi une translocation inférieure aux autres protéines. Il semble que les résidus chargés négativement pourraient interférer dans une certaine proportion, sur l'efficacité de la translocation. L'anticorps polyclonal  $\alpha$ G2, dirigé contre la  $\gamma$ -zéine (Ludevid et al., 1985, Plant Sci. 41, 41-48) ne pouvant être utilisé pour distinguer entre la  $\gamma$ -zéine de type

sauvage et les  $\gamma$ -zéines modifiées, un anticorps  $\alpha$ PL dirigé contre un peptide synthétique contenant la séquence d'acides aminés riche en lysine a été préparé. On a ensuite testé si des protéines modifiées synthétisées in-vitro étaient reconnues à la fois par les anticorps  $\alpha$ G2 et les anticorps  $\alpha$ PL. Cette expérience est illustrée à la figure 4B et la figure 4C où les transcrits synthétiques de la  $\gamma$ -zéine, P20 $\gamma$ Z, H30 $\gamma$ Z, H45 $\gamma$ Z et N13 $\gamma$ Z étaient traduits in-vitro et dans lesquels les produits de traduction étaient immunoprécipités avec  $\alpha$ PL (Fig. 4B) et avec  $\alpha$ G2 (Fig. 4C). Ces résultats indiquent que les  $\gamma$ -zéines riches en lysine étaient reconnues par les deux anticorps (voir Fig. 4B et C, lignes 2 à 5) et que la  $\gamma$ -zéine était seulement reconnue par le sérum  $\alpha$ G2 (Fig. 4B et C, ligne 1). Ainsi la spécificité des anticorps  $\alpha$ PL pour les protéines modifiées a permis de distinguer la  $\gamma$ -zéine riche en lysine de la  $\gamma$ -zéine endogène lorsque les gènes modifiés étaient exprimés dans des cellules d'endosperme. Prises ensemble ces expériences ont montré que la présence de séquences riches en lysine ne perturbait pas la fonction de translocation à travers la membrane ou le comportement immunologique de la  $\gamma$ -zéine.

#### Analyse de l'expression des $\gamma$ -zéines enrichies en lysine dans les endospermes de maïs

Pour explorer si la  $\gamma$ -zéine modifiée riche en lysine était exprimée et accumulée dans les cellules de l'endosperme, des graines de stade 17 DAP ont été bombardées avec l'ADN contenant les séquences codant pour la protéine des quatre constructions (Fig. 3) sous le contrôle du promoteur CaMV (P20 $\gamma$ Z et H20 $\gamma$ Z) et du promoteur de la  $\gamma$ -zéine (H45 $\gamma$ Z et N13 $\gamma$ Z). Des constructions comprenant des promoteurs anti-sens ou dépourvues de promoteurs ont été utilisées comme témoins. Après 24 heures de transformation des endospermes, les protéines totales ont été extraites et l'expression des  $\gamma$ -zéines modifiées a été testée par immunoblotting en

utilisant l'anticorps  $\alpha$ PL. La figure 6A montre que les gènes chimères de  $\gamma$ -zéine contenant l'insertion (Pro-Lys)<sub>n</sub> après le domaine Pro-X (P20 $\gamma$ Z) ou le remplaçant (H45 $\gamma$ Z et H30 $\gamma$ Z) étaient fortement exprimés et les produits de traduction accumulés efficacement dans les cellules de l'endosperme (Fig. 6A, lignes 3, 4 et 5). Pour chaque ligne, les extraits de protéine correspon-  
 5 daient à 1/3 environ d'un endosperme bombardé, ce qui a permis d'estimer que la quantité de protéines modifiées P20 $\gamma$ Z, H30 $\gamma$ Z et H45 $\gamma$ Z par endosperme atteignait le niveau du nanogramme. De plus, aucune différence quantitative entre les niveaux d'expression des gènes chimères  
 10 sous le contrôle du promoteur CaMV et du promoteur de la  $\gamma$ -zéine n'a été observée, confirmant les résultats ci-dessus décrits obtenus avec les protéines marqueurs GUS et LUC (Fig. 5). On a noté que l'anticorps  $\alpha$ PL reconnaissait une protéine d'environ 30 kD présente dans les extraits de protéine totale, même dans les endospermes non transformés (voir la  
 15 bande faible présente dans les quatre lignes de la figure 6A). Une procédure d'extraction séquentielle de protéine a permis de constater que cette protéine n'était pas une protéine de réserve soluble dans un milieu aqueux.

Comme le montre la figure 6A (ligne 2) aucune trace de la protéine  
 20 N13 $\gamma$ Z n'a été détectée, indiquant que le gène chimère correspondant n'était pas exprimé dans les cellules de l'endosperme ou que la protéine N13 $\gamma$ Z était dégradée. Des ARNs des endospermes transformés avec les ADNs codant pour les protéines H45 $\gamma$ Z et N13 $\gamma$ Z ainsi que les ARNs des endospermes non transformés ont été analysés. A partir des ARNs totaux,  
 25 les ADNc ont été préparés et amplifiés par PCR en utilisant des amorces spécifiques. La Fig. 6B montre l'analyse en Southern blot de trois échantillons d'ADNc hybridés avec un oligonucléotide codant pour une séquence K(Pro-Lys)<sub>4</sub> utilisée comme sonde. Les résultats indiquent que le gène N13 $\gamma$ Z était correctement exprimé (Fig. 6B, ligne 3). La présence de  
 30 bandes dans les échantillons H45 $\gamma$ Z et N13 $\gamma$ Z, mais pas dans les

endospermes non transformés, a suggéré que la protéine N13 $\gamma$ Z était dégradée pendant les 24 heures de l'incubation. A partir de ces observations, les inventeurs ont conclu que le site d'insertion des séquences riches en lysine était critique pour la stabilité de la  $\gamma$ -zéine modifiée.

#### La $\gamma$ -zéine enrichie en lysine est accumulée dans le corps protéiques

En dehors de la teneur en lysine des séquences Pro-X, les protéines P20 $\gamma$ Z, H30 $\gamma$ Z et H45 $\gamma$ Z avaient des caractéristiques communes avec la  $\gamma$ -zéine de type sauvage : le peptide signal, le domaine répété en tandem N-terminal et la région riche en cystéine C-terminale étaient apparentées. Il a semblé important de déterminer si ces domaines restaient totalement fonctionnels préservant le ciblage et la formation des corps protéiques ou si les séquences riches en lysine créaient un environnement spécial dans lequel ces propriétés pouvaient être perturbées. Pour tester ceci, on a investigué si les  $\gamma$ -zéines modifiées étaient capables de s'accumuler dans les corps protéiques. Un fractionnement sub-cellulaire a été réalisé avec les endospermes transformés. Les homogénats d'endospermes bombardés ont été chargés sur des gradients de sucrose discontinus (20 %, 50 % et 70 % de sucrose) et toutes les fractions collectées ont été analysées par immunoblotting. P20 $\gamma$ Z, H30 $\gamma$ Z et H45 $\gamma$ Z sédimentés sur la fraction des corps protéiques et aucune quantité significative de ces protéines n'a été détectée, soit dans le surnageant, soit dans la fraction microsomale (Fig. 7A, lignes 2, 2 et 3). Bien que dans les expériences faites préalablement *in-vitro* (Fig. 4A) aient permis de constater que les  $\gamma$ -zéines modifiées nouvellement synthétisées subissaient une translocation dans les microsomes canins, on a testé ici si les protéines modifiées exprimées *in-vivo* dans des cellules d'endosperme, subissaient une translocation dans la membrane du réticulum

endoplasmique et restaient à l'intérieur des corps protéiques dérivés du réticulum endoplasmique. Pour cette raison, des corps protéiques isolés ont été soumis à une digestion par la protéinase K dans des tampons isotoniques (contenant 20 % de sucrose) ou après un choc osmotique dans l'eau (Fig. 7B). Les protéines protégées contre la dégradation protéolytique par des enzymes peuvent être entourées par une membrane et un traitement par des solutions détergeantes ou hypotoniques a résulté dans la digestion des protéines (Walter et Blobel, 1983, Method Enzymol. 96, 84-93). Une comparaison des intensités de bandes, après digestion par la protéinase K dans des milieux comprenant du sucrose ou de l'eau, a permis de révéler que les protéines P20 $\gamma$ Z, H30 $\gamma$ Z et H45 $\gamma$ Z étaient protégées de la digestion dans des tampons isotoniques (lignes 1, 3 et 5) mais étaient partiellement digérées dans l'eau (lignes 2, 4 et 6).

L'expression des gènes modifiés de la  $\gamma$ -zéine dans les cellules de la couche du sub-aleurone de l'endosperme par bombardement biolistique a permis d'observer que des  $\gamma$ -zéines riches en lysines étaient largement accumulées à l'exception du cas où les séquences riches en lysine étaient positionnées 5 résidus en amont de l'extrémité C-terminale du polypeptide de la  $\gamma$ -zéine. A partir de cette expression et d'études immunocytochimiques sur des corps protéiques isolés, les inventeurs ont démontré que des  $\gamma$ -zéines riches en lysine sont correctement accumulées dans ces organites et sont co-localisées avec les protéines endogènes  $\gamma$ -zéine et  $\alpha$ -zéine.

Des corps protéiques isolés à partir des endospermes P20 $\gamma$ Z ont été examinés par marquage de type immunogold et microscopie électronique. Sur des sections ultra fines incubées avec l'anticorps  $\alpha$ P-L (Fig. 8A), le marquage à l'or a été détecté à l'intérieur des corps protéiques, indiquant que la protéine P20 $\gamma$ Z riche en lysine était accumulée à l'intérieur de ces organites. Dans des sections incubées seulement avec l'anticorps  $\alpha$ P-L, l'immunomarquage avait lieu uniquement sur quelques corps protéiques

(contenant la  $\gamma$ -zéine riche en lysine), la plus grande partie des corps protéiques isolés n'était pas immunomarkée avec des anticorps  $\alpha$ P-L parce qu'ils correspondaient à des cellules d'endosperme non transformées. Pour déterminer si la  $\gamma$ -zéine riche en lysine était co-localisée avec les polypeptides  $\alpha$ -zéines et  $\gamma$ -zéines, un double marquage en immuno-électromicroscopie a été réalisé sur des corps protéiques isolés en utilisant les anticorps  $\alpha$ Z et  $\alpha$ P-L (Fig. 8B) et  $\alpha$ G2 et  $\alpha$ P-L (Fig. 8C et D). La figure 8B montre une micrographie de la section transversale de deux corps protéiques marqués avec l'anticorps  $\alpha$ P-L (particule d'or de 15 nm) et avec l'anticorps  $\alpha$ Z (particule d'or de 5 nm). Le résultat de l'immunocoloration a montré que la protéine P20 $\gamma$ Z est accumulée dans les corps protéiques et co-localisée avec l' $\alpha$ -zéine (voir l'étendue du marquage de l' $\alpha$ -zéine sur toute la surface du corps protéique). Par ailleurs, des sections tangentielles (Fig. 8B, voir les flèches) et transversales (Fig. 8D) des corps protéiques ont été incubées avec l'anticorps  $\alpha$ P-L (particule d'or 15 nm) et avec l'anticorps  $\alpha$ G2 (particule d'or de 5 nm). Dans les deux cas, la protéine P20 $\gamma$ Z était co-localisée avec les polypeptides  $\gamma$ -zéines. On a pu noter que les sections tangentielles du corps protéique (Fig. 8A, C, voir les flèches) étaient facilement distinguées des sections transversales du corps protéique dans la mesure où les premiers présentaient une densité plus forte en électrons et que le marquage de la  $\gamma$ -zéine était étendu sur toute la section. Au contraire les sections transversales présentaient une densité inférieure et le marquage de la  $\gamma$ -zéine était localisé sur la membrane entourant le corps protéique. Dans les deux cas, la localisation du marquage de la  $\gamma$ -zéine riche en lysine suivait celle de la  $\gamma$ -zéine endogène.

**-B) Préparation de plantes génétiquement modifiées, exprimant des  $\gamma$ -zéines riches en lysine**

1) **Obtention et utilisation de cal de maïs comme cible pour la transformation génétique.**

5           La transformation génétique du maïs, quelle que soit la méthode employée (electroporation; Agrobacterium, microfibres, canon à particules) requiert généralement l'utilisation de cellules indifférenciées en divisions rapides ayant conservé une aptitude à la régénération de plantes entières. Ce type de cellules compose le cal friable embryogène (dit de type II) de  
10   maïs.

          Ces cals sont obtenus à partir d'embryons immatures de génotype Hi II ou (A188 x B73) selon la méthode et sur les milieux décrits par Armstrong (Maize Handbook ; (1994) M. Freeling, V. Walbot Eds ; pp 665-671). Les cals ainsi obtenus sont multipliés et maintenus par repiquages  
15   successifs tous les quinze jours sur le milieu d'initiation.

          Des plantules sont ensuite régénérées à partir de ces cals en modifiant l'équilibre hormonal et osmotique des cellules selon la méthode décrite par Vain et al. (Plant Cell Tissue and Organ Culture (1989 18 : 143-151). Ces plantes sont ensuite acclimatées en serre où elles peuvent être  
20   croisées ou autofécondées.

2) **Utilisation du canon à particules pour la transformation génétique du maïs.**

25           Le paragraphe précédent décrit l'obtention et la régénération des lignées cellulaires nécessaires à la transformation ; on décrit ici une méthode de transformation génétique conduisant à l'intégration stable des gènes modifiés dans le génome de la plante. Cette méthode repose sur l'utilisation d'un canon à particules identique à celui décrit par J. Finer  
30   (Plant Cell Report (1992) 11 :323-328) ; les cellules cibles sont des fragments des cals décrits dans le paragraphe 1. Ces fragments d'une

surface de 10 à 20 mm<sup>2</sup> ont été disposés, 4h avant bombardement, à raison de 16 fragments par boîte au centre d'une boîte de petri contenant un milieu de culture identique au milieu d'initiation, additionné de 0,2 M de mannitol + 0,2M de sorbitol. Les plasmides portant les gènes à introduire sont purifiés sur colonne Qiagen® en suivant les instructions du fabricant. Ils sont ensuite précipités sur des particules de tungsten (M10) en suivant le protocole décrit par Klein (Nature (1987) 327 :70-73). Les particules ainsi enrobées sont projetées vers les cellules cibles à l'aide du canon et selon le protocole décrits par J. Finer (Plant Cell Report (1992) 11 : 323-328).

Les boîtes de cals ainsi bombardées sont ensuite scellées à l'aide de Scellofrais® puis cultivées à l'obscurité à 27°C. Le premier repiquage a lieu 24h après, puis tous les quinze jours pendant 3 mois sur milieu identique au milieu d'initiation additionné d'un agent sélectif dont la nature et la concentration peuvent varier selon le gène utilisé (voir paragraphe 3). Les agents sélectifs utilisables consistent généralement en composés actifs de certains herbicides (Basta®, Round up®) ou certains antibiotiques (Hygromycine, Kanamycine...).

On obtient après 3 mois ou parfois plus tôt, des cals dont la croissance n'est pas inhibée par l'agent sélectif, habituellement et majoritairement composés de cellules résultant de la division d'une cellule ayant intégré dans son patrimoine génétique une ou plusieurs copies du gène de sélection. La fréquence d'obtention de tels cals est d'environ 0,8 cal par boîte bombardée.

Ces cals sont identifiés, individualisés, amplifiés puis cultivés de façon à régénérer des plantules (cf. paragraphe 1). Afin d'éviter toute interférence avec des cellules non transformées toutes ces opérations sont menées sur des milieux de culture contenant l'agent sélectif.

Les plantes ainsi régénérées sont acclimatées puis cultivées en serre où elles peuvent être croisées ou autofécondées.

### 3) Utilisation du gène Bar pour la production plantes de maïs génétiquement modifiées ayant incorporé et exprimant le gène H45yZ

Le gène Bar de *Streptomyces hygroscopicus* code pour une phosphinotricine acyl transférase (PAT) qui inactive par acétylation la phosphinotricine-molécule active de l'herbicide Basta®. Les cellules portant ce gène sont donc rendues résistantes à cet herbicide et peuvent être sélectionnées par son intermédiaire. Pour la transformation des céréales la séquence codante du gène Bar est sous le contrôle d'une région régulatrice permettant une expression forte et constitutive dans les cellules végétales. Une telle région est avantageusement constituée par le promoteur et le premier intron du gène Actine de riz tels que décrits par Mc Elroy (Mol. Gen. Genet. (1991) 231 : 150-160).

Ce gène chimérique est cloné sur un plasmide permettant son amplification par *Escherichia coli*. Ce plasmide (pDM 302 Cao (Plant Cell Report (1992) 11:586-591) après amplification puis purification sur colonne Qiagen® peut être utilisé en transformation génétique du maïs en utilisant par exemple la méthode décrite dans l'exemple précédent. Dans ce cas les milieux de culture destinés à la sélection des cellules transformées sont additionnés de 2 mg/L de Phosphinotricine.

Pour l'introduction du gène H45yZ une technique dite de cotransformation est avantageusement utilisée : les gènes de sélection (Bar) et d'intérêt (H45yZ) sont portés par des plasmides-indépendants. Dans le cas de l'utilisation du canon à particules on procède à une coprécipitation des plasmides sur les particules de tungstène, la quantité totale d'ADN précipité sur les particules restant identique à ce qu'elle est dans le protocole standard (5µg d'ADN pour 2,5 mg de particules) chaque plasmide représentant environ la moitié du poids total d'ADN utilisé.

L'expérience montre qu'avec cette méthode, la cointégration des plasmides dans les cellules végétales est l'évènement le plus fréquent (de

l'ordre de 90%) c'est-à-dire que pratiquement chaque plante ayant intégré le gène Bar et sélectionnée par son intermédiaire portera aussi le gène H45yZ. Le niveau de coexpression (pourcentage de plantes sélectionnées exprimant le gène H45yZ) est habituellement de l'ordre de 70%.

5 Les gènes ainsi introduits sont généralement liés au sens génétique, le gène H45yZ peut ainsi avantageusement être suivi dans les descendance grâce à la résistance à l'herbicide qui lui est étroitement associée.

10 La quantité de protéine modifiée est déterminée par les méthodes décrites dans l'exemple A, notamment par immunoblotting sur extraits protéiques de grains de maïs immatures ou matures, prélevés en pool sur des plantes résistantes au BASTA.

15 **4) Exemple explicitant l'étape d'introduction des transgènes et notamment du gène codant pour la H45yZ, pour la modification du phénotype opaque 2 du maïs**

Amélioration de maïs opaque 2 par l'introgression du gène de  $\gamma$ -zéine enrichie en lysine.

20 Les plantes transformées décrites dans l'exemple précédent, présentent à la fois, une résistance au Basta et l'expression d'une  $\gamma$ -zéine enrichie en lysine sont utilisées. Leur pollen est utilisé pour féconder les plantes de maïs opaque 2 de la lignée W64Ao2 qui ne contient qu'un seul gène  $\gamma$ -zéine. Cette lignée a été obtenue auprès du Maize Stock Center.

25 Les plantes de ces descendance F1 sont sélectionnées pour leur résistance au Basta et ensuite autofécondées. Les graines F2 ainsi produites sont analysées pour le phénotype opaque sur une table lumineuse et les grains opaque ou les grains vitreux sont semés et évalués pour la résistance au Basta. Dans le cas où tous les grains opaques sont

30 sensibles au Basta, la démonstration est faite que l'introduction du gène  $\gamma$ -

zéine enrichi en lysine dans la plante considérée, complémente le phénotype opaque 2.

Dans un deuxième temps, dans ces plantes Basta résistantes, des individus de génotype o2/o2 et ne possédant qu'un seul gène  $\gamma$ -zéine sur le chromosome 7 sont sélectionnés à l'aide des sondes moléculaires codant pour le gène Opaque 2 et pour la  $\gamma$ -zéine. Ces dernières révélant des fragments de restriction polymorphes et seuls les individus présentant le patron type de la lignée W64o2 sont retenus (Lopes M.A. et al, 1995, Mol. Gen. Genet. 19 : 247, 603-613).

Ces individus ont un contenu en lysine qui est en moyenne équivalent à celui du maïs o2, voire supérieur. A partir de ces individus, toute introgression dans des variétés ELITE, du caractère « forte teneur en lysine » est observée via la détermination de la résistance au BASTA et de la présence de l'allèle o2 détectée par RFLP.

#### 5) Expression de $\gamma$ -zéines enrichies en lysine dans *Arabidopsis thaliana*.

Pour obtenir une transformation stable, les constructions plasmidiques P20 $\gamma$ Z et pH30 $\gamma$ Z clonées dans le plasmide Bluescript KS (-) ont été insérées sous forme de fragments HincII/XbaI dans le vecteur binaire pBin19 (Bevan, M. Nucl. Acids Res. 12: 8711-8721 (1984)), contenant le promoteur 35S du virus de la mosaïque du chou-fleur (CaMV) et les signaux de formation de l'extrémité 3' et de polyadénylation du gène de l'octopine synthétase (ocs). Les nouveaux plasmides ont été appelés p19P20 $\gamma$ Z et p19H30 $\gamma$ Z (figure 12).

Les vecteurs binaires contenant les séquences codant pour les protéines P20 $\gamma$ Z et H30 $\gamma$ Z (p19P20 $\gamma$ Z et p19H30 $\gamma$ Z) ont été transférés dans la souche LBA4404 d'*Agrobacterium tumefaciens*. Des plantes *Arabidopsis* ecotype RLD ont été transformées selon la méthode décrite par Valvekens, D., Van Montagu, M. et Van Lijsebettens, M. ((1988) Proc.

Natl.Acad.Sci.USA 85: 5536-5540). Pour chaque construction, 10 plantes transgéniques ont été criblées par analyse en immunoblot en utilisant un antisérum obtenu contre la  $\gamma$ -zéine ( $\alpha$ G2, Ludevid et al. 1985). Les plantes contenant les taux les plus élevés (correspondant à environ 0,1% de la quantité totale de protéines présentes dans les feuilles d'*Arabidopsis*) de produits transgéniques dans la génération F1, ont été choisies pour obtenir la génération F2. Ces plantes ont également été sélectionnées pour l'expression de la protéine recherchée.

Des plantes transgéniques entières, sélectionnées dans un milieu contenant de la kanamycine ont été homogénéisées dans de l'azote liquide. Les protéines transgéniques ont été extraites sélectivement avec une solution contenant de l'éthanol/ de l'acide chlorhydrique HCl 0,125 N dans une proportion de 3:1 (v/v), avec 5% de mercaptoéthanol et d'inhibiteurs de protéase. Les protéines extraites avec cette solution ont été précipitées dans 5 volumes d'acétone et analysées par SDS-PAGE et immunoblotting. Les extraits protéiques des plantes non transgéniques ont été utilisés en tant que contrôles. Les protéines résultant de l'insertion des séquences K(P-K)<sub>4</sub> dans la  $\gamma$ -zéine ont été exprimées correctement dans les plantes *Arabidopsis thaliana* en utilisant le promoteur constitutif 35.S du CaMV. Sur les immunoblots, les anticorps  $\alpha$ G2 et  $\alpha$ PL reconnaissaient des bandes d'électrophorèse correspondant aux protéines P20 $\gamma$ Z et H30 $\gamma$ Z. Ces bandes migraient avec des masses moléculaires apparentes conformes à celles que l'on avait préalablement observées dans les expériences de traduction / translocation *in vitro* (30 kD et 26 kD respectivement). Comme on l'a observé dans les plantes transgéniques *Arabidopsis* exprimant la  $\gamma$ -zéine (Geli et al. Plant Cell 6: 1911-1922 (1994)), les protéines P20 $\gamma$ Z et H30 $\gamma$ Z migraient sous la forme de deux bandes d'électrophorèse à savoir les bandes correspondant à 36 et 30 kD pour P20 $\gamma$ Z et les bandes correspondant à 32 et 26 kD pour H30 $\gamma$ Z. Les bandes supérieures pourraient correspondre à des produits ayant subi des

modifications post-traductionnelles. Une telle modification post-traductionnelle n'a pas été détectée dans les endospermes de maïs transformés. Ce résultat suggère que la modification apparaîtrait lorsque ces protéines sont exprimées dans un système hétérologue tel que

5 Arabidopsis thaliana.

## REVENDICATIONS

1. Oligonucléotide comprenant au moins un enchaînement  
 5 codant pour un polypeptide répondant à la formule  $(P-K)_n$ , dans laquelle :
  - n est un entier supérieur ou égal à 2,
  - P représente un résidu d'acide aminé Proline,
  - K représente un résidu d'acide aminé Lysine,
  - le signe "-" symbolise une liaison entre les deux résidus d'acides
  - 10 aminés et notamment une liaison de type peptidique, les n unités (P-K) étant liées entre elles également par de telles liaisons par exemple de type peptidique.
2. Oligonucléotide selon la revendication 1, comprenant un enchaînement codant pour un polypeptide répondant à la formule  $(P-K)_n$ ,  
 15 dans laquelle n est un entier supérieur ou égal à 3, de préférence n est égal à 4, 5, 6, 7, 8, 9, 10 ou 15.
3. Oligonucléotide selon la revendication 1 ou la revendication 2, comprenant un enchaînement codant pour un polypeptide répondant à la formule  $(P-K)_n$ , dans laquelle la séquence des n unités (P-K) est interrompue par un ou plusieurs résidus d'acides aminés différents  
 20 des résidus P ou K.
4. Oligonucléotide selon l'une quelconque des revendications 1 à 3, caractérisé en ce que l'enchaînement codant pour le polypeptide comprenant les n unités (P-K) est complété à son extrémité 5' et/ou à son  
 25 extrémité 3', par un ou plusieurs codons, codant par exemple pour au moins un résidu lysine à l'extrémité N-terminale du polypeptide formé.
5. Oligonucléotide selon la revendication 4, caractérisé en ce qu'il code pour un polypeptide répondant à la formule  $(P-K)$ , à la formule  $-K(P-K)_4$  ou à la formule  $2K(P-K)_4$ .

6. Séquence de nucléotides recombinant comprenant un enchaînement de nucléotides codant pour une protéine de réserve d'une plante, caractérisée en ce qu'elle comprend en outre un oligonucléotide selon l'une quelconque des revendications 1 à 5, inséré en un site de l'enchaînement de nucléotides choisi de façon telle que :

- l'expression de la séquence de nucléotides dans une cellule végétale déterminée permet d'obtenir une protéine de réserve **modifiée** localisée de façon identique ou similaire à la protéine de réserve **normale** qui serait exprimée dans les mêmes conditions dans la même cellule par l'enchaînement de nucléotides codant correspondant, et/ou

- la protéine de réserve modifiée codée par la séquence de nucléotides recombinant est reconnue immunologiquement par des anticorps produits contre la protéine de réserve normale correspondante.

7. Séquence de nucléotides selon la revendication 6, caractérisée en ce que l'enchaînement de nucléotides codant qu'elle comprend, code pour une protéine de réserve naturellement pauvre en lysine.

8. Séquence de nucléotides selon la revendication 7, caractérisée en ce que l'enchaînement de nucléotides codant qu'elle comprend, code pour une protéine de réserve naturellement produite par une plante utilisable en alimentation animale ou humaine.

9. Séquence de nucléotides selon la revendication 7 ou la revendication 8, caractérisée en ce que l'enchaînement de nucléotides codant qu'elle comprend, code pour une protéine de réserve naturellement produite par une plante de la famille des céréales.

10. Séquence de nucléotides selon la revendication 7 ou la revendication 8, caractérisée en ce que l'enchaînement de nucléotides codant qu'elle comprend, code pour une protéine de réserve naturellement produite par une plante de la famille des légumineuses ou des crucifères.

11. Séquence de nucléotides selon la revendication 9, caractérisée en ce que l'enchaînement de nucléotides codant qu'elle comprend, code pour une protéine de réserve de maïs.

5 12. Séquence de nucléotides selon la revendication 11, caractérisée en ce que l'enchaînement de nucléotides codant qu'elle comprend, code pour une protéine de réserve de la famille des zéines.

13. Séquence de nucléotides selon la revendication 12, caractérisée en ce que l'enchaînement de nucléotides codant qu'elle comprend, code pour une protéine de réserve est la  $\gamma$ -zéine du maïs.

10 14. Séquence de nucléotides selon la revendication 13, caractérisée en ce que l'enchaînement de nucléotides codant pour la  $\gamma$ -zéine de maïs qu'elle contient, répond à la séquence présentée sur la Figure 9.

15 15. Séquence de nucléotides selon la revendication 7 ou la revendication 8, caractérisée en ce que l'enchaînement de nucléotides codant qu'elle comprend, code pour une protéine de réserve d'une plante choisie parmi les suivantes : le soja, le tournesol, le tabac, le blé, l'avoine, la luzerne, le riz, le colza, Arabidopsis thaliana.

20 16. Séquence de nucléotides selon la revendication 13 ou la revendication 14, caractérisée en ce que l'oligonucléotide est inséré à la place ou à la suite du domaine Pro-X ou dans le domaine Pro-X naturellement présent dans la  $\gamma$ -zéine du maïs.

25 17. Séquence de nucléotides recombinante, caractérisée en ce qu'elle comprend une séquence de nucléotides selon l'une quelconque des revendications 1 à 16 sous le contrôle d'un promoteur d'expression.

30 18. Séquence de nucléotides recombinante selon la revendication 17, caractérisée en ce que le promoteur est un promoteur spécifique d'un tissu cellulaire déterminé, par exemple un promoteur spécifique pour l'expression dans les graines, et/ou dans les feuilles de plantes.

19. Séquence de nucléotides selon la revendication 17, caractérisée en ce que le promoteur d'expression est celui de la  $\gamma$ -zéine du maïs.

20. Séquence de nucléotides selon la revendication 17, caractérisée en ce que le promoteur d'expression est le promoteur CaMV35S.

21. Séquence de nucléotides selon l'une quelconque des revendications 16 à 20, caractérisée en ce qu'elle code pour l'un des polypeptides P20 $\gamma$ Z ou PH45 $\gamma$ Z répondant respectivement aux séquences représentées aux figures 11 et 10.

22. Vecteur de clonage et/ou d'expression, caractérisé en ce qu'il comprend, en un site non essentiel pour sa réplication, une séquence de nucléotides selon l'une quelconque des revendications 1 à 20.

23. Vecteur de clonage et/ou d'expression, caractérisé en ce qu'il s'agit de l'un des plasmides pP20 $\gamma$ Z (CNCM N° I-1640), pH30 $\gamma$ Z ou pH45 $\gamma$ Z (CNCM N° I-1639).

24. Polypeptide codé par une séquence selon l'une quelconque des revendications 1 à 17.

25.  $\gamma$ -zéine de maïs modifiée riche en lysine, caractérisée en ce qu'elle est codée par la séquence de nucléotides selon la revendication 13 ou 14.

26.  $\gamma$ -zéine de maïs modifiée enrichie en lysine, caractérisée en ce que sa séquence d'acides aminés est modifiée par au moins un polypeptide répondant à la formule (P-K) $_n$  ou à la formule K-(P-K) $_n$ , dans laquelle :

- n est un entier supérieur ou égal à 2,
- P représente un résidu d'acide aminé Proline,
- K représente un résidu d'acide aminé Lysine,
- le signe "-" symbolise une liaison entre les deux résidus d'acides aminés et notamment une liaison de type peptidique, les n unités (P-K)

étant liées par des liaisons notamment de type peptidique, ledit polypeptide répondant à la formule  $(P-K)_n$  ou  $K-(P-K)_n$  étant substitué à une séquence naturellement présente dans la  $\gamma$ -zéine de maïs normale ou étant inséré avec délétion d'un ou plusieurs acides aminés de la séquence d'acides aminés de la  $\gamma$ -zéine de maïs normale, ou étant ajouté dans la séquence d'acides aminés de la  $\gamma$ -zéine normale, le site d'insertion du polypeptide étant choisi de façon telle que :

- lorsque la  $\gamma$ -zéine modifiée riche en lysine est produite dans une cellule hôte, notamment une cellule végétale, elle est localisée de façon identique ou similaire à la  $\gamma$ -zéine de maïs normale qui serait produite dans les mêmes conditions, dans la même cellule hôte, et/ou,

- la  $\gamma$ -zéine de maïs modifiée est reconnue par des anticorps dirigés contre la  $\gamma$ -zéine normale de maïs.

**27.**  $\gamma$ -zéine de maïs modifiée selon la revendication 26, caractérisée en ce qu'il s'agit de la protéine P20 $\gamma$ Z ou de la protéine H30 $\gamma$ Z ou de la protéine H45 $\gamma$ Z.

**28.** Cellule hôte recombinante, caractérisée en ce qu'elle comprend une séquence de nucléotides selon l'une quelconque des revendications 1 à 17.

**29.** Cellule hôte selon la revendication 28, caractérisée en ce qu'il s'agit d'une bactérie, par exemple de E. coli ou Agrobacterium tumefaciens.

**30.** Cellule hôte selon la revendication 28, caractérisée en ce qu'il s'agit d'une cellule végétale.

**31.** Cellule hôte selon la revendication 30, caractérisée en ce qu'il s'agit d'une cellule de semences de plante.

**32.** Cellule hôte selon la revendication 31, caractérisée en ce qu'il s'agit d'une cellule d'endosperme de semences de maïs.

33. Cellule hôte selon la revendication 32, caractérisée en ce qu'elle contient, intégrée de façon stable dans son génome, une séquence de nucléotides selon la revendication 13.

34. Cellule hôte selon la revendication 32, caractérisée en ce qu'elle produit une  $\gamma$ -zéine de maïs modifiée riche en lysine selon la revendication 26 ou la revendication 27.

35. Cellule hôte selon la revendication 30, caractérisée en ce qu'il s'agit d'une cellule de soja, de tournesol, de tabac, de blé, d'avoine, de luzerne, de riz, de colza, d'Arabidopsis.

36. Semences produisant un polypeptide selon l'une quelconque des revendications 24 à 27.

37. Plante produisant un polypeptide selon l'une quelconque des revendications 24 à 27.

38. Plante selon la revendication 37, caractérisée en ce qu'il s'agit du maïs.

39. Semences obtenues à partir des plantes selon la revendication 37 ou 38.

40. Procédé de production de plantes ou de semences exprimant une protéine de réserve modifiée, caractérisé en ce qu'il comprend les étapes de :

a) transformation d'une cellule végétale, avec une séquence de nucléotides selon l'une quelconque des revendications 6 à 17, ou un vecteur selon l'une quelconque des revendications 22 ou 23, dans des conditions permettant l'expression de façon stable et fonctionnelle de la protéine de réserve modifiée codée par la séquence de nucléotides;

b) régénération de plantes à partir de la cellule de plante transformée de l'étape a), pour obtenir des plantes exprimant la protéine de réserve modifiée,

c) le cas échéant, obtention de semences à partir des plantes modifiées obtenues à l'étape b).

**41.** Procédé selon la revendication 40, caractérisé en ce que la plante est le maïs et la protéine de réserve est la  $\gamma$ -zéine.

pP20<sub>γ</sub>Z

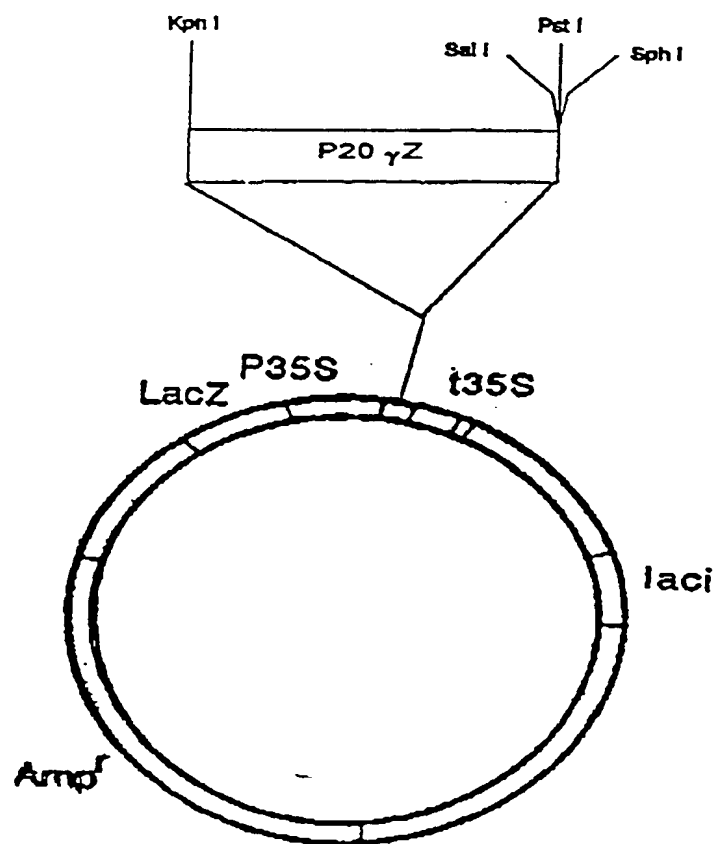


FIGURE 1

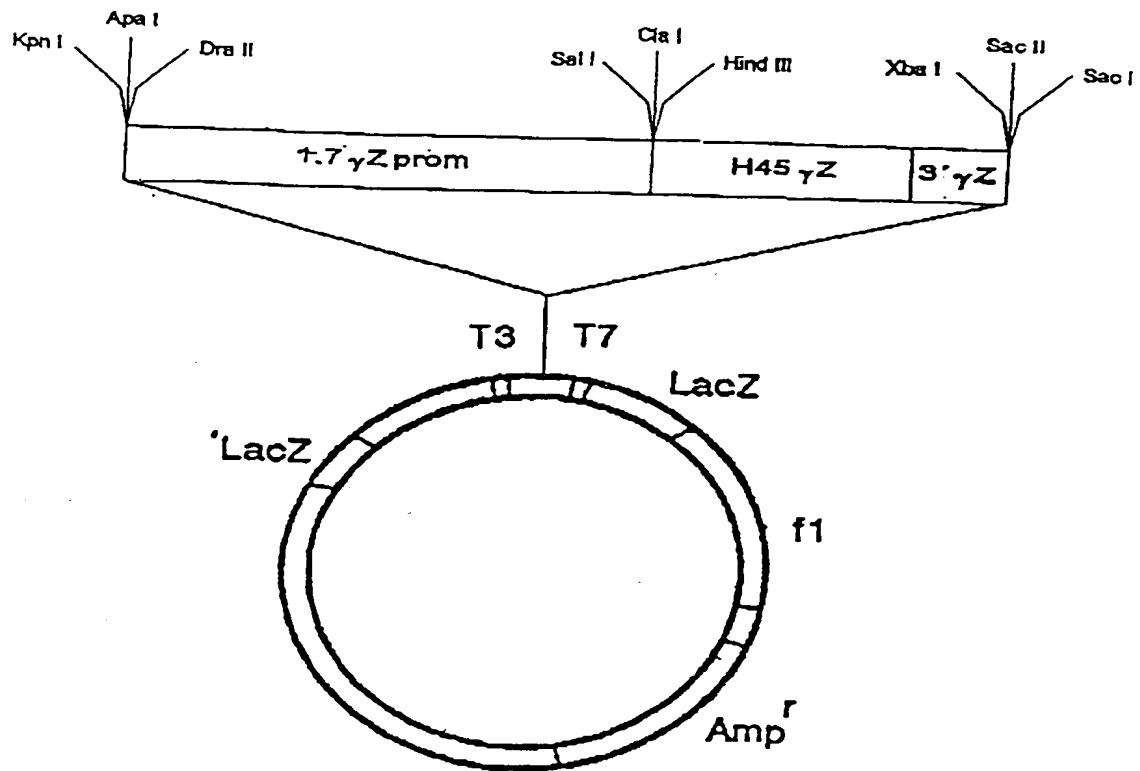
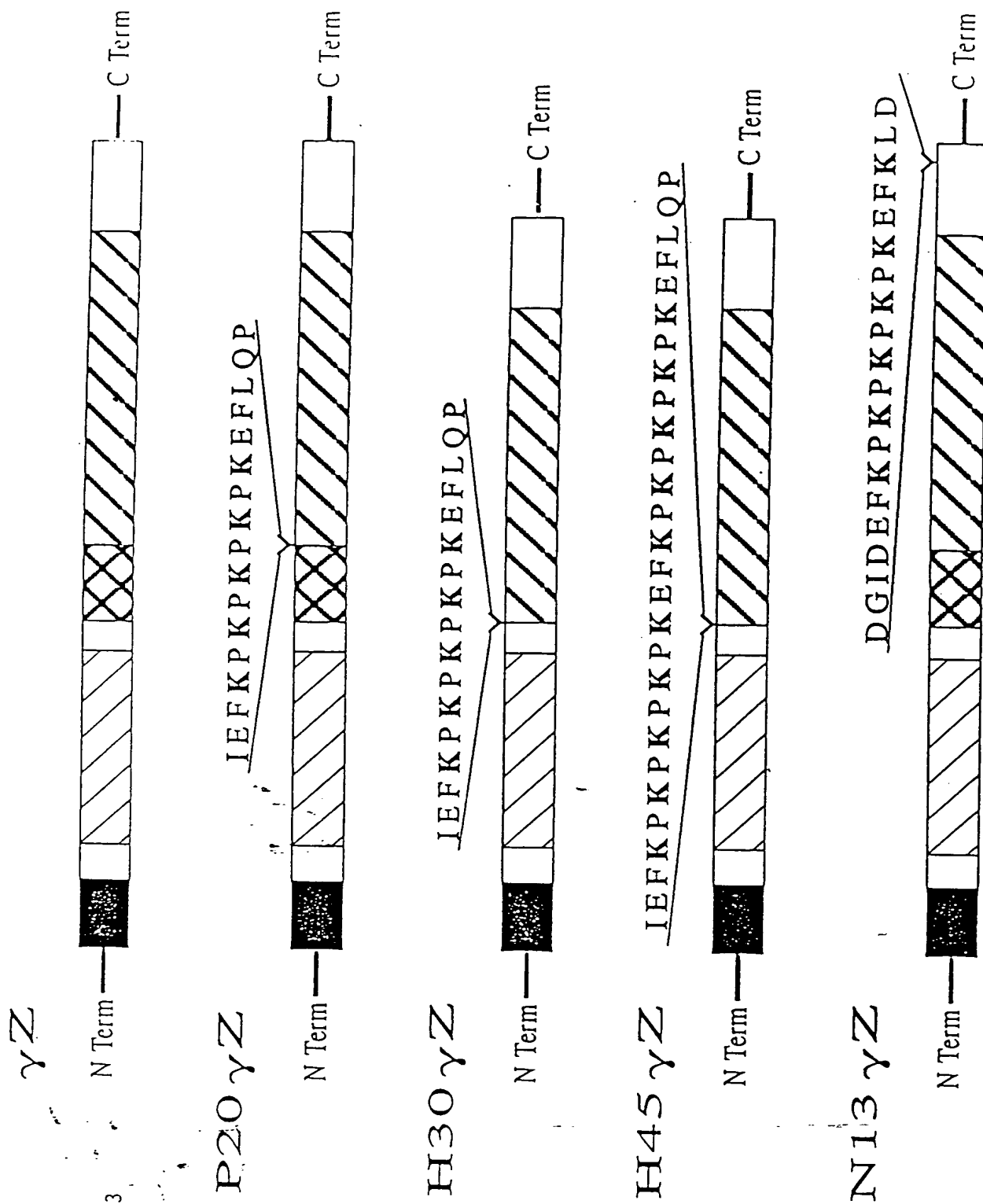
pH45<sub>γ</sub>Z

FIGURE 2



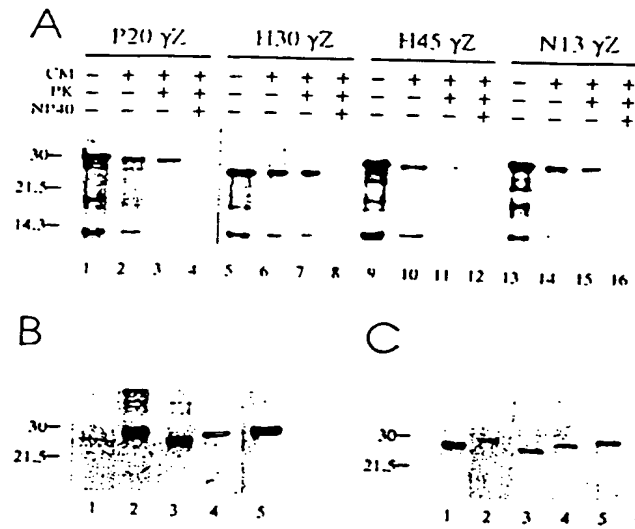


FIGURE 4

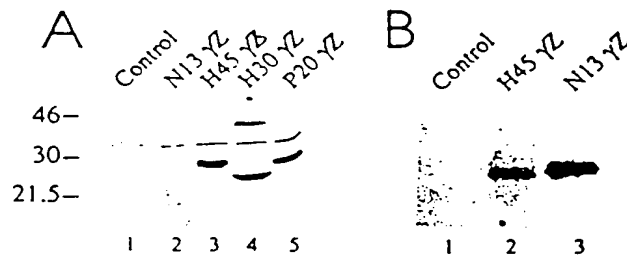


FIGURE 6

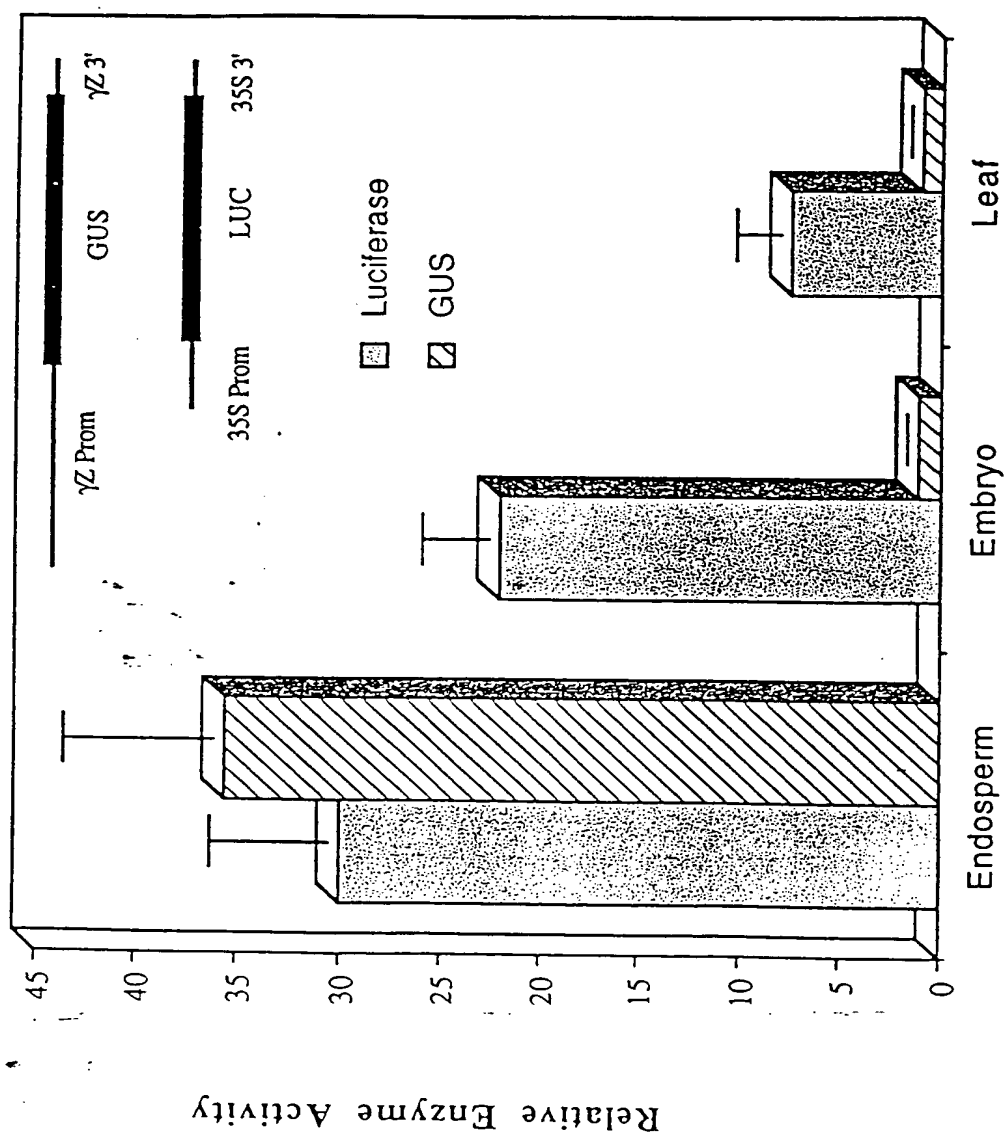


FIGURE 5

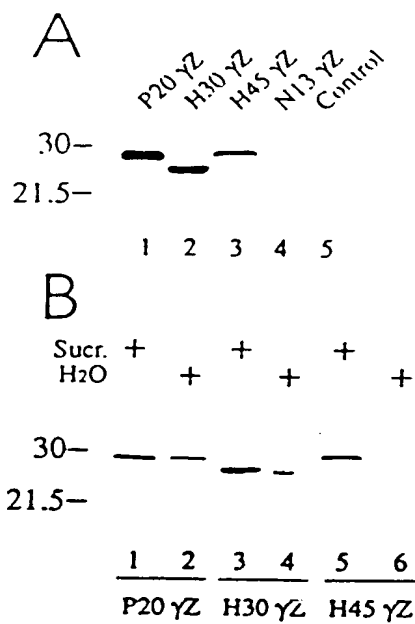


FIGURE 7

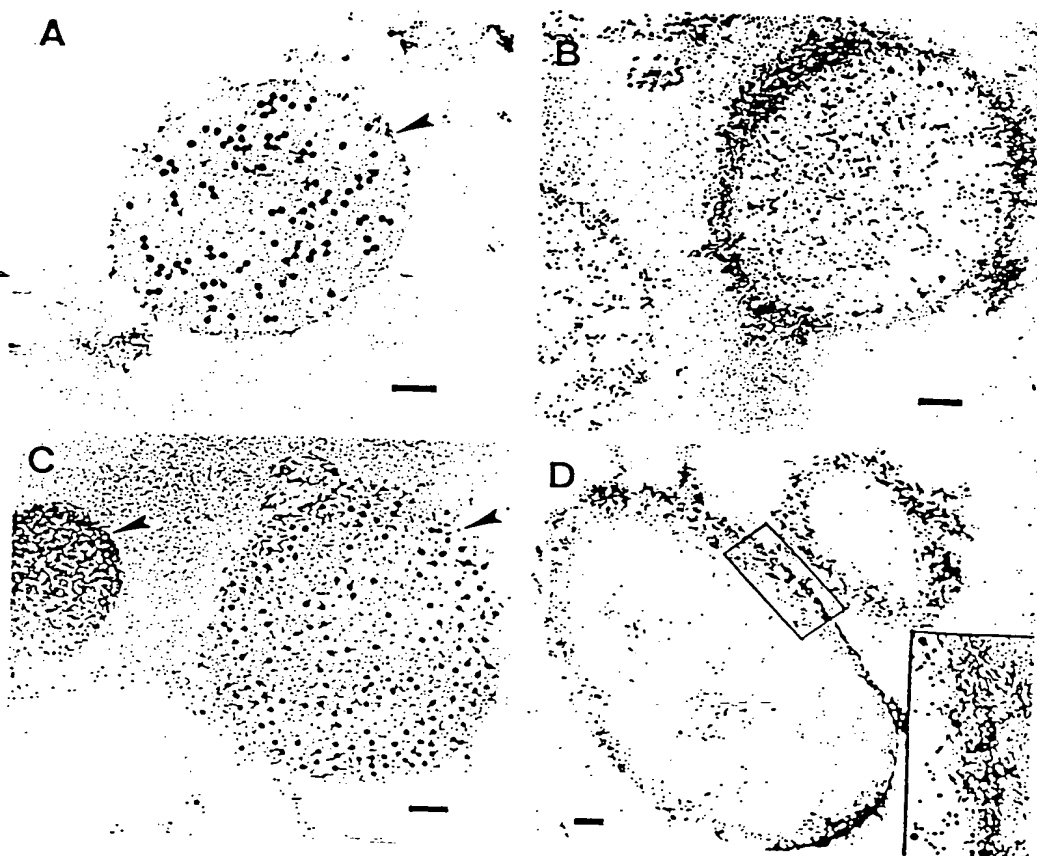


FIGURE 8

**$\gamma$ -Zeina**

10	20	30	40	50	60
ATGAGGGTGT	TGCTCGTTG	CCCTCGCTCT	CCTGGCTCTC	GCTGCGAGCG	CCACCTCCACG
M R V L L V	A L A L L A	L A A S A T	S T		
70	80	90	100	110	120
CATACAAGCG	GGCGGCTG	CGGCTGCCAG	CCACCGCCG	CGCGGTTTCAT	CTACCGCCG
H T S G G C	G C Q P P P	P V H L P P	P P V		
130	140	150	160	170	180
CATCTGCCAC	CTCCGGTTC	ACCTGCCAC	CTCCGGTGC	ATCTCCACCG	CCGGTCCACCTG
H L P P P V	H L P P P V	H L P P P V	H L P P P V	H L	
190	200	210	220	230	240
CCGCCGCGCG	GTCCACCTG	CCACCGCGG	TCCATGTGCC	GCGCGGTTCA	TCTGCCGCGCG
P P P V H L	P P P V H V	P P P V H L	P P		
250	260	270	280	290	300
CCACCATGCC	ACTACCCTA	CTCAACCGC	CCCCGGCCT	CAGCCTCATC	CCCCAGCCAC
P P C H Y P	T Q P P R P	Q P H P Q P	H P Q P H P		
310	320	330	340	350	360
TGCCCGTGCC	AACAGCCGC	ATCCAAGCC	CGTGCCAGCT	GCAAGGAAC	CTGCGGCGTT
C P C Q Q P	H P S P C Q	L Q G T C G	V G		
370	380	390	400	410	420
AGCACCCCGA	TCCTGGGCC	AGTGCGTCA	GAGTTTCTG	AGGCATCAGT	GCAGCCCGAC
S T P I L G	Q C V E F L	R H Q C S P	T A		
430	440	450	460	470	480
ACGCCCTACT	GCTCGCCTC	AGTGCCAGT	CGTTGCGGC	CAGCAGTGT	TGCCAGCAGCT
T P Y C S P	Q C Q S L R	Q Q C C Q Q	L R		
490	500	510	520	530	540
CAGGTGGAGC	CGCAGCAC	CGGTACCAG	GCGATCTTC	CGGCTTGGT	CCTCCAGTCC
Q V E P Q H	R Y Q A I F	G L V L Q S	I L		
550	560	570	580	590	600
CAGCAGCAGC	CGCAAAGC	GGCCAGGT	CGCGGGGCT	GTTGGCGGC	GCAGATAGCG
Q Q Q P Q S	G Q V A G L	L A A Q I A	Q Q		

FIGURE 9

610 620 630 640 650 660  
CTGACGGCGATGTGCGGCCTGCAGCAGCCGACTCCATGCCCTACGCTGCTGCCGGCGGT  
L T A M C G L Q Q P T P C P Y A A A G G  
670  
GTCCCCCACTGA  
V P H -

FIGURE 9 (Suite)

## H45 yZ

10	20	30	40	50	60
ATGAGGGTGT	TGCTCGTTG	CCCTCGCTCT	CCTGGCTCTC	GCTGCGAGCG	CCACCTCCACG
M R V L L V	A L A L L A	L A A S A T	S T		
70	80	90	100	110	120
CATACAAGCG	GGCGGTGCGG	CTGCCAGCCAC	CGCCGCGGTT	CATCTACCGCC	GCGCGGTG
H T S G G C	G C Q P P P	P V H L P P	P P V		
130	140	150	160	170	180
CATCTGCCAC	CTCCGGTTCAC	CTGCCACCTCC	GGTGATCTCCC	ACCGCCGGTCC	ACCTG
H L P P P V	H L P P P V	H L P P P V	H L		
190	200	210	220	230	240
CCGCCGCGGT	TCCACCTGCC	ACCGCCGGTCC	ATGTGCCGCG	CGCGGTTCAT	CTGCCGCCG
P P P V H L	P P P V H V	P P P V H L	P P		
250	260	270	280	290	300
CCACCATGCC	ACTACCCTACT	CAACCGCCCC	GGATCGAATT	CAAACCAAAG	CCAAAGCCG
P P C H Y P	T Q P P R I	E F K P K P	K P		
310	320	330	340	350	360
AAGCCAAAAG	AATTCAAACCA	AAGCCAAAGC	CGAAGCCAAA	AGAATTCTGC	CAGCCCCCTG
K P K E F K	P K P K P K	P K E F L Q	P L		
370	380	390	400	410	420
CAGGGAACCT	TGCGGCGTTG	GCGAGCACCC	CGATCCTGGG	CCAGTGCCTC	GAGTTTCTGAGG
Q G T C G V	G S T P I L	G Q C V E F	L R		
430	440	450	460	470	480
CATCAGTGC	CAGCCCGACG	GCGACGCCCT	ACTGCTCGCT	CAGTGCCAGT	CGTTGCGGCAG
H Q C S P T	A T P Y C S	P Q C Q S L	R Q		
490	500	510	520	530	540
CAGTGTGCC	CAGCAGCTC	AGGCAGGTG	GAGCCGCAGC	ACCGGTACC	CAGGCGATCTTCGGC
Q C C Q Q L	R Q V E P Q	H R Y Q A I	F G		
550	560	570	580	590	600
TTGGTCCTC	CAGTCCATC	CTGCAGCAGC	AGCCGCAAAG	CGGCCAGGT	CGCGGGGCTGTG
L V L Q S I	L Q Q P Q S	G Q V A G L	L		
610	620	630	640	650	660
GCGGCGCAG	ATAGCGCAG	CAACTGACG	GCGATGTG	CGGCCCTGC	CAGCAGCCGACTCCATGC
A A Q I A Q	Q L T A M C	G L Q Q P T	P C		

FIGURE 10

670 680 690  
| | |  
CCCTACGCTGCTGCCGGCGGTGTCCCCCACTGA  
P Y A A A G G V P H -

FIGURE 10 (Suite)

## P20 γZ

10	20	30	40	50	60
ATGAGGGTGTGCTCGTTGCCCTCGCTCTCCTGGCTCTCGCTGCGAGCGCCACCTCCACG					
M R V L L V A L A L L A A S A T S T					
70	80	90	100	110	120
CATAACAAGCGGCGGCTGCGGCTGCCAGCCACCGCCGCGCGTTTCATCTACCGCCGCGCGGTG					
H T S G G C G C Q P P P P V H L P P P V					
130	140	150	160	170	180
CATCTGCCACCTCCGGTTTCACCTGCCACCTCCGGTGTCATCTCCCACCGCCGGTCCACCTG					
H L P P P V H L P P P V H L P P P V H L					
190	200	210	220	230	240
CCGCCGCGCGGTCCACCTGCCACCGCCGGTCCATGTGCCGCGCGCGGTTCATCTGCCGCGG					
P P P V H L P P P V H V P P P V H L P P					
250	260	270	280	290	300
CCACCATGCCACTACCCTACTCAACCGCCCCGGCCTCAGCCTCATCCCCAGCCACACCCA					
P P C H Y P T Q P P R P Q P H P Q P H P					
310	320	330	340	350	360
TGCCCGTGCCAACAGCCGCATCCAAGCCCGTGCCAGATCGAATTCAAACCAAAGCCAAAG					
C P C Q Q P H P S P C Q I E F K P K P K					
370	380	390	400	410	420
CCGAAGCCAAAGAATTCTGTCAGCCCCTGTCAGGGAACCTGCGGCGTTGGCAGCACCCCG					
P K P K E F L Q P L Q G T C G V G S T P					
430	440	450	460	470	480
ATCCTGGGCCAGTGCCTCGAGTTTCTGAGGCATCAGTGCAGCCCGACGGCGACGCCCTAC					
I L G Q C V E F L R H Q C S P T A T P Y					
490	500	510	520	530	540
TGCTCGCCTCAGTGCCAGTCGTTGCCGCGCAGTGTGTCAGCAGCTCAGGCAGGTGGAG					
C S P Q C Q S L R Q Q C C Q Q L R Q V E					
550	560	570	580	590	600
CCGCAGCACCGGTACCAGGCGATCTTCGGCTTGGTCCTCCAGTCCATCCTGCAGCAGCAG					
P Q H R Y Q A I F G L V L Q S I L Q Q Q					
610	620	630	640	650	660
CCGCAAAGCGGCCAGGTGCGGGGCTGTTGGCGGCGCAGATAGCGCAGCAACTGACGGCG					

FIGURE 11

P Q S G Q V A G L L A A Q I A Q Q L T A

670 680 690 700 710 720

ATGTGCGGCCTGCAGCAGCCGACTCCATGCCCCCTACGCTGCTGCGCGCGGTGTCCCCCACC

M C G L Q Q P T P C P Y A A A G G V P H

TGA

FIGURE 11 (Suite)

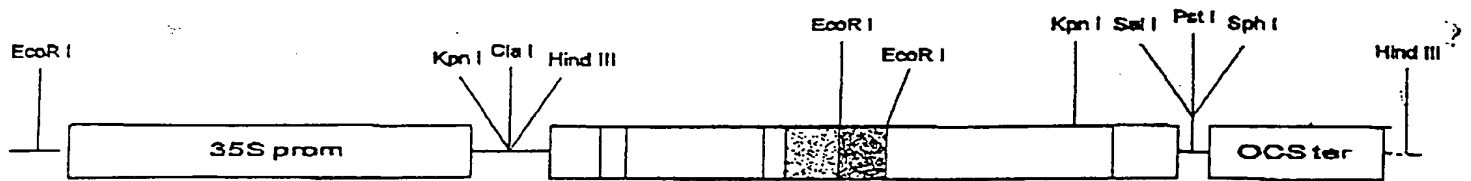
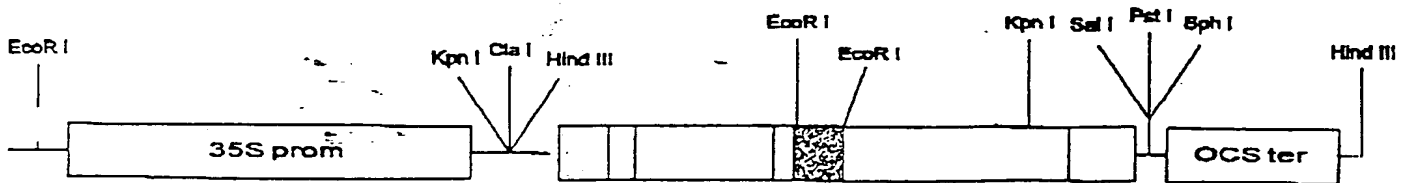
pBin 19 P20  $\gamma$ ZpBin 19 H30  $\gamma$ Z

FIGURE 12

**THIS PAGE BLANK (USPTO)**